Novel Approaches to Thalassaemia Gene Therapy

Azevedo Ferreira, Sónia

Awarding institution:
King's College London

NOVEL APPROACHES TO THALASSAEMIA GENE THERAPY

by;

Sonia Marisa Azevedo Ferreira

King's College London

2012

Department of Medical and Molecular Genetics

Guys Hospital

University of London

DECLARATION

I hereby declare that the work presented in this thesis is mine, unless otherwise stated.

ACKNOWLEDGEMENTS

My first word has necessarily to be to my first supervisor, Dr. Michael Antoniou. Thank you for allowing me to do a PhD in your group, for your guidance, support, patience and criticism. Secondly, I would like to thank my second supervisor Dr. Simon Waddington and Dr. Suzy Buckley for their enthusiasm and much appreciated help with all animal work!

I would like to acknowledge the Nuclear Biology/ Gene Therapy and Cell Expression lab, for their continual and much appreciated support.

Finally, a few words to my loved ones; although no words could possibly describe how grateful I am to them. I would like to thank my parents for their support and my little sis, Diana, for always being present although physically miles away. Ultimately, I have a special word for Billy, whom very patiently humored my temper and moods in the last few months, for the strength, support, and much appreciated care.

TABLE OF CONTENTS

CHAPTER TWO - MATERIALS AND METHODS

CHAPTER FIVE - DEVELOPMENT OF IN UTERO GENE THERAPY APPROACHES FOR INHERITED DISEASES – I 225

ABREVIATIONS

α	alpha
ºC	degrees Celsius
β	beta
βLCR	human beta globin locus control region
AAV	adeno-associated viruses
Amp	ampicillin
BM	bone marrow
BMT	bone marrow transplant
bp	base pair
BSA	bovine serum albumin
cDNA	copy deoxyribonucleic acid
CFTR	cystic fibrosis transmembrane conductance regulator
cPPT	central polypurine tract
CMV	cytomegalovirus
CpG	cytosine/guanine DNA doublet
cPPT	central polypurine tract
CV	coefficient of variegation
cpm	counts per minute
DES	human desmin
dH_2O	distilled water
DMEM	Dubecco's modified eagle medium
DNA	deoxyribonucleic acid
dNTPs	dinucleotide triphosphate
dpc	days postconception
dsDNA	double-stranded deoxyribonucleic acid
eGFP	enhanced green fluorescent protein
EDTA	ethylenediaminetetraacetic acid

ELISA	enzyme linked imunosorbent assay
Env	envelope
FACS	fluorescent analysis cell sorting
FCS	foetal calf serum
γ	gamma
g	gram
GAPDH	glyceraldehde-3-phosphate dehydrogenase
Gag	group specific antigens
h	hour
HSC	hematopoietic stem cell
HbA	adult haemoglobin
HBB	human β-eta globin gene
HBA	human α-globin chain
HBB	human β-globin chain
HBG	human γ-globin chain
HBE	human ε-globin chain
HEL	human erythroleukemia lineage
HIV-1	human immunodeficiency virus- type 1
HLA	human leukocyte antigen
HMGA2	high-mobility group AT-hook 2
HS1-HS5	core sites from the HBB locus control region
HS	hypersensitivite sites
HSC	haematopoietic stem cell
HSCT	haematopoietic stem cell transplantation
ID	integration deficient
In	integrase
IP	integration proficient
kb	kilo base pairs
KO	knocked out
LB	Luria Bertani

kDa	kilo dalton
LCR	locus control region
LTR	long terminal repeat
LV	lentiviral vector
LVs	lentiviral vectors
µg	microgram
µl	microlitre
M	molar
m	milli (10^{-3})
µ	microliter (10^{-6})
MEL	murine erythroleukemia lineage
MHC	major histocompability complex
mg	milligram
min	minutes
ml	millilitre
MA	matrix
MFI	mean fluorescence intensity
MHC	major histocompatibility complex
MLV	murine leukaemia virus
MOI	multiplicity of infection
mRNA	messager RNA
OD	optical density
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEV	positional effect variegation
PIC	pre-integration complex
PPT	polypurine tract
qPCR	quantitative polymerase chain reaction
RNA	ribonucleic acid
RRE	*rev* responsive element

rpm	revolutions per minute
RT	reverse transcriptase
SCID	Severe combined immunodeficiency
SCA	sickle cell anaemia
SIN	self-inactivating lentiviral vector
SFFV	spleen forming focus virus
S/MAR	scaffold matrix attachment region
Ss	single stranded
T cell	thymus derived lymphocyte
TAE	tris-acetate-EDTA
TU/ml	transfection units per ml
UCOE	ubiquitously-acting open chromatin element
VSV-G	vesicular stomatitis virus-glycoprotein G
v/v	volume per volume
WPRE	woodchuck post-transcriptional regulatory element
w/w	weight per weight
ψ	packaging signal

Chapter One

NOVEL APPROACHES TO THALASSAEMIA GENE THERAPY - INTRODUCTION

1.1 Introduction to Haemoglobinopathies

1.1.1 The Haemoglobin molecule

Haemoglobin is an iron-containing oxygen transport protein responsible for the delivery of oxygen into the cells. As **Figure 1.1** illustrates, haemoglobin is a tetramer protein complex formed of two types of globin sub-units each bound to a haeme group - two β-globin chains (HBB) and two α-globin chains (HBA) in adult haemoglobin (Madigan & Malik, *2006*).

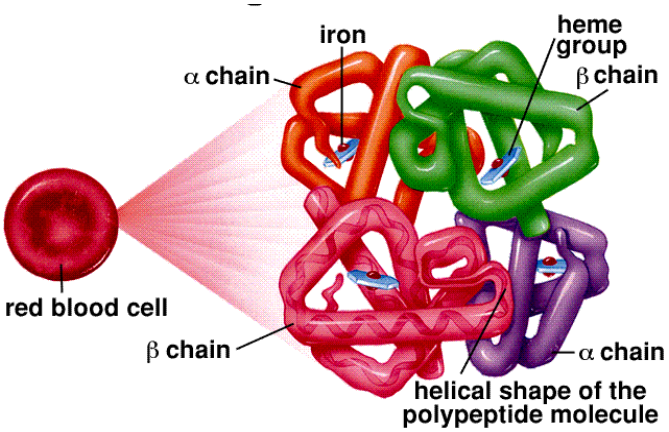


Figure 1.1 - Structure of human haemoglobin molecule. The α and β subunits, and the iron-containing haeme groups in blue are shown. Each erythrocyte contains approximately 270 million haemoglobin molecules (*image from http://gassama.myweb.uga.edu/).*

The name haemoglobin comes from *haeme* and *globin*, since each subunit of haemoglobin is a globular protein with an embedded haeme (or haem) group. Each haeme group contains an iron atom responsible for the binding of oxygen.

The haemoglobinopathies are a group of blood disorders characterised by the dysfunction of haemoglobin. Dysfunctions of the haemoglobin protein are caused by the imbalance of the α and β polypeptide ratio (as in the thalassaemias) or due to abnormalities in the function of the HBB haemoglobin chain caused by mutations

resulting in structural variants that impair haemoglobin normal mechanism of action (as in sickle cell anaemia).

1.1.2. α and β Thalassaemias

The thalassaemias are mostly autosomal recessive diseases, being the most common monogenic disorders in the world. Studies estimate 270 million carriers worldwide, of which 80 million are of β-thalassaemia trait carrier status. Depending on the chain imbalance in the haemoglobin tetramer molecule, several types of α and β can occur.

1.1.2.1 α-thalassaemia

Alpha-thalassaemia is the generic term that encompasses all of those conditions in which there is a deficit in the production of HBA chains, and typically results from deletions involving *HBA1* and *HBA2* at their chromosome 16 loci. Underproduction of HBA gives rise to excess HBB-type chains which form HBG tetramers *in utero* and HBB tetramers post-natally (Harteveld & Higgs, *2010*).

Carriers are normally unaffected and/or show minimal anaemia.

There are two types of α thalassaemia that can cause serious health problems. The first is *HbH* disease, which can vary in severity depending on whether the patient has 3, 2 or 1 remaining functional *HGA*. Affected individuals are mainly compound heterozygotes (with some homozygous patients having been reported) for α-thalassaemia and have a moderately severe anaemia characterised by the presence of excess HBB in the peripheral red blood cells. Severe post-natal α-thalassaemia major or HbH disease is characterised by severe anaemia, heart ailments, hepatosplenomegaly or enlargement of liver and spleen, and urinary or genital abnormally, jaundice, bone deformities particularly in the upper jaw and forehead and constant exhaustion. Recent medical advances have allowed survival through *in*

utero-blood transfusion. However, lifelong blood transfusions and constant medical care are necessary to sustain the child's life (*http://www.alphathalassemia.net*).

The second most severe type of α-thalassaemia is known as haemoglobin *Bart hydrops fetalis syndrome* or *Hb Barts syndrome*. Affected individuals synthesise very little or no α globin chains and have a very severe form of anaemia which causes death *in utero* or neonatally (Harteveld & Higgs, 2010). In the presence of little or no HGA chains excess HGG chains during foetal development form HGG tetramers, which bind oxygen with very high affinity and do not readily release this to tissues. The resulting severe hypoxia gives rise to *hydrops fetalis*; that is, death before birth.

1.1.2.2 β-thalassaemia

Beta-thalassaemia in the majority of cases is inherited as a recessive disorder. However, dominant mutations have also been reported. The genetic defect of β-thalassaemia arises from the absence or low levels of HBB chains resulting in the precipitation of un-tetramerised α-chains, causing ineffective erythropoiesis *(*Sadelain *et al.*, 2005) and consequent damage to red blood cells, which become vulnerable to mechanical injury and lyse easily. The resulting anemia and ineffective erythropoiesis that is associated with β-thalassaemia may be severe and is accompanied by bone marrow expansion and extramedullary haematopoiesis in the liver, spleen and other sites such as paravertebral masses (Rund & Rachmilewitz, 2005). In the most severe forms, found in homozygote or compound heterozygote individuals, the anaemia is lethal within the first few years of life in the absence of any treatment (Sadelain *et al.*, 2005).

The β-thalassaemias are grouped into 3 categories: thalassaemia *major*, thalassaemia *intermedia* and thalassaemia *minor*. Most heterozygotes (carriers) are clinically normal and largely unaware of their genetic condition. Patients with thalassaemia intermedia are mostly homozygotes or compound heterozygotes. Inheritance of two copies of β-thalassaemia causing genes, however, results in life-threatening anemia and transfusion dependence for survival.

Although deletions of *HBB* are uncommon, more than 200 mutations have been reported to date that give rise to β-thalassaemia; the large majority are point mutations in functionally important regions of *HBB* (Galanello & Origa, 2010). Many mutations leading to β-thalassaemia have been informative with regard to the sequences and molecular mechanisms that control globin gene switching (*see section **1.2.3** of this report*).

A list of common mutations according to severity and ethnic distribution is listed in **Table 1.1**.

Population	**β-gene mutation**	**Severity**
Indian	-619 del	β^0
Mediterranean	-101 C→T	β^0
Black	-88 C→T	β^0
Mediterranean; African	-87 C→G	β^0
Japanese	-31 A→G	β^{++}
African	-29 A→G	β^{++}
Southeast Asian	-28 A→C	β^{++}
Mediterranean; Asian Indian	IVS1-nt1 G→A	β^0
East Asian; Asian Indian	IVS1-nt5 G→C	β^0
Mediterranean	IVS1-nt6 T→C	$\beta^{+/++}$
Mediterranean	IVS1-nt110 G→A	β^+
Chinese	IVS2-nt654 C→T	β^+
Mediterranean	IVS2-nt745 C→G	β^+
Mediterranean	codon 39 C→T	β^0
Mediterranean	codon 5 –CT	β^0
Mediterranean; African-American	codon 6 -A	β^0
Southeast Asian	codon 41/42 -TTCT	β^0
African-American	AATAAA to AACAAA	β^{++}
Mediterranean	AATAAA to AATGAA	β^{++}
Mediterranean	codon 27 G→T Hb (Hb Knossos)	β^{++}
Southeast Asian	codon 79 G>A (Hb E)	β^{++}
Malaysia	Codon 19 G>A (Hb Malay)	β^{++}

Table 1.1 – Most common types of β-thalassaemia - ethnic distribution and severity of different β-gene mutations found so far. (Data as tabulated by Galanello & Origa, 2010).

1.1.3 Current therapeutic options

1.1.3.1 Regular blood transfusions

The treatment for severe forms of thalassaemia (*major* and some *intermedia*) consists of regular packed red blood cell (RBC) transfusion therapy every 2-4 weeks. Transfusions are life-saving and aim to correct the anaemia, suppress the massive erythropoiesis, and inhibit increased gastrointestinal absorption of iron (Sadelain *et al.*, 2005). Nevertheless, transfusion therapy invariably causes iron accumulation, and because humans naturally have no physiological excretory mechanism to deal with iron overload, patients maintained on a regular transfusion regime progressively develop clinical manifestations of iron accumulation: hypogonadism (35-55% of the patients), hypothyroidism (9-11%), hypoparathyroidism (4%), diabetes (6-10%), liver fibrosis, and heart dysfunction (33%) (Cunningham *et al.*, 2004).

The complications of iron overload, together with the sequelae of the anaemia, the ineffective erythropoiesis, and the chelation therapy itself, are the major causes of morbidity and mortality leading to an increased toxicity in the endocrine system and liver, pancrease damage and cardiomyopathy in transfusion dependent β-thalassaemias (Quek & Thein, 2007). It is therefore imperative to conduct a parallel treatment of iron chelation and management of secondary complications of iron overload. Iron chelators form a complex with the excess iron in the tissues, and facilitate its excretion from the body. Chelation therapies reduce tissue iron stores, prevent iron-induced organ damage, and reduce morbidity and mortality. Regular treatment consists of slow subcutaneous injection of desferrioxamine, which must be administered for 12h per day, 6-days per week; a treatment that is time consuming, expensive, and also very painful for the patient. Patients also risk acquiring viral infections transmitted with blood transfusions (Cunningham *et al*., 2004).

Scientific advances in recent years saw the introduction of two orally administered iron chelators, namely deferiprone (*Feriprox*) and deferasivox (*Exjade*) have provided a new hope for thalassaemic patients. These new chelators bring the improved quality of life by reducing or eliminating the use of daily subcutaneous

pump administration. In addition, oral deferiprone has been proved to be effective in the removal of myocardiac iron (Meerpohl *et al.*, 2010).

1.1.3.2 Bone marrow transplant

Allogeneic haematopoietic stem cell transplant (HSCT) from bone marrow (BMT) or umbilical cord blood (UCBT) is conceptually the simplest and the only approach so far that may lead to a definitive cure for β-thalassaemia. Patients who benefit the most from this treatment are those with early but severe transfusion-dependent disease and have little co-morbidity since HSCT success rates are lower for older patients (~30%) or those who have reduced overall health status caused by inadequate chelation. A major complication associated with this approach is the low probabilities of finding a leukocyte antigen (HLA)-matching donor, preferably a sibling, to minimise the risks of graft rejection and graft-versus-host-disease (GVHD) (Sadelain *et al*., 2005).

BMT from unrelated donors has also been performed on a more limited scale with individuals with β-thalassaemia and provided that selection of the donor is based on stringent criteria of HLA compatibility, and that individuals have limited iron overload, results are comparable to those obtained when the donor is a compatible sibling. Unfortunately, to date there is still a limited number of cases benefitting from this approach - there are very few healthy volunteers willing to be BM donors, and therefore the chances of finding an HLA matched donor in the BM cell bank are limited (Gaziev *et al., 2010).*

1.1.3.3 Umbilical Cord blood stem cell transplantation

A good alternative for couples who already have a child with β-thalassaemia is Umbilical Cord blood (UCB) transplantation from a related donor. UCB transplants are currently performed at a 10 times lower ratio than BM transplants, as the lower amount of transplanted HSCs is clinically reflected in a greater incidence of engrave failure and prolonged time to engraft. This risk is however, significantly offset by lower rates of acute GVHD given the lower number and mostly naïve T cells in the transplanted population (Stanevsky *et al.*, 2009).

The search for and HLA matching donor remains, although with less restriction when compared with BMT and the search is made easier with the existence of Public and private UCB cell banks.

1.1.4 Global distribution of the pathology

Thalassaemia has a high incidence in a broad band extending from the Mediterranean basin and parts of Africa, throughout the Middle East, the Indian sub-continent, South-East Asia, Melanesia and into the Pacific Islands. The carrier frequency for β-thalassaemia in these areas ranges from 1% to 20%, but rarely greater, while that for the milder forms of αthalassaemia is much higher, ranging from 10-20% in parts of sub-Saharan Africa, through to 40% or more in some Middle Eastern and Indian populations, to as high as 80% in northern Papua New Guinea and isolated groups in north-east India. The α^o-thalassaemias are more restricted in their distribution, occurring at high frequencies only in parts of South-East Asia and the Mediterranean basin and therefore pose less of a global health problem that the β thalassaemias (Weatherall & Clegg, 2001). The geographic distribution of the thalassaemias is shown in **Figure 1.2**.

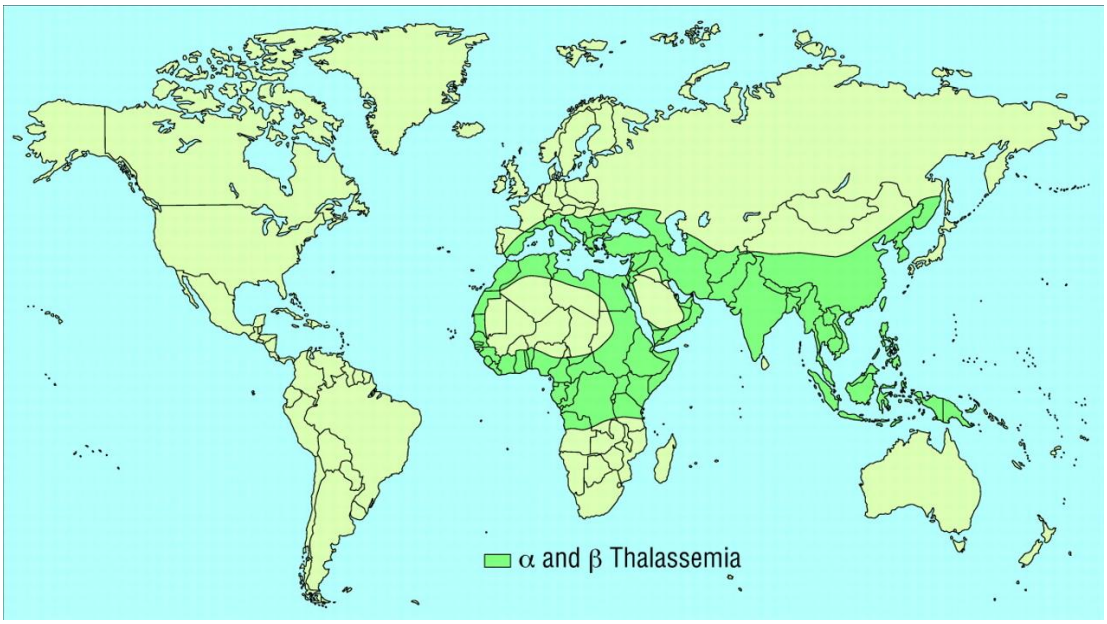


Figure 1.2 - Global distribution of α and β thalassaemia. The world health organization (WHO) estimates that, on a global scale, at least 2.9% of adults are carriers for thalassaemia (Weatherall & Clegg, 2001).

It is thought that all of these disorders have been selected because they protect carriers from the ravages of malaria and epidemiological evidence supports this hypothesis (Weatherall, 2010).

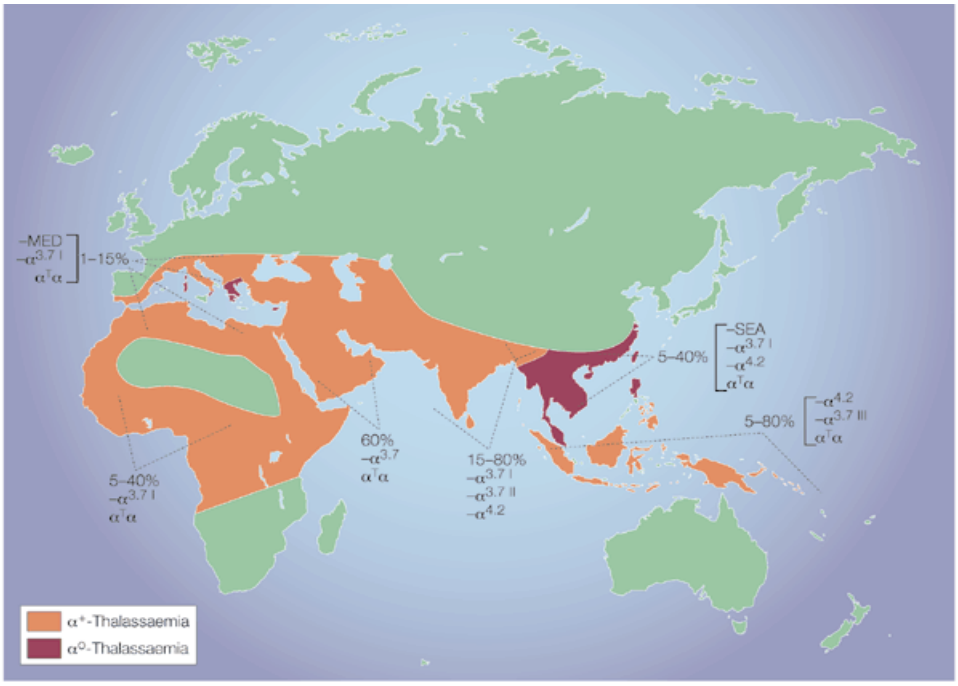


Figure 1.3 - Global α-thalassaemia distribution. α^{+} is more frequent in the general population whereas α^{o} is more prevalent in Greece and south east asia (Image taken from Weatherall D.J., 2001).

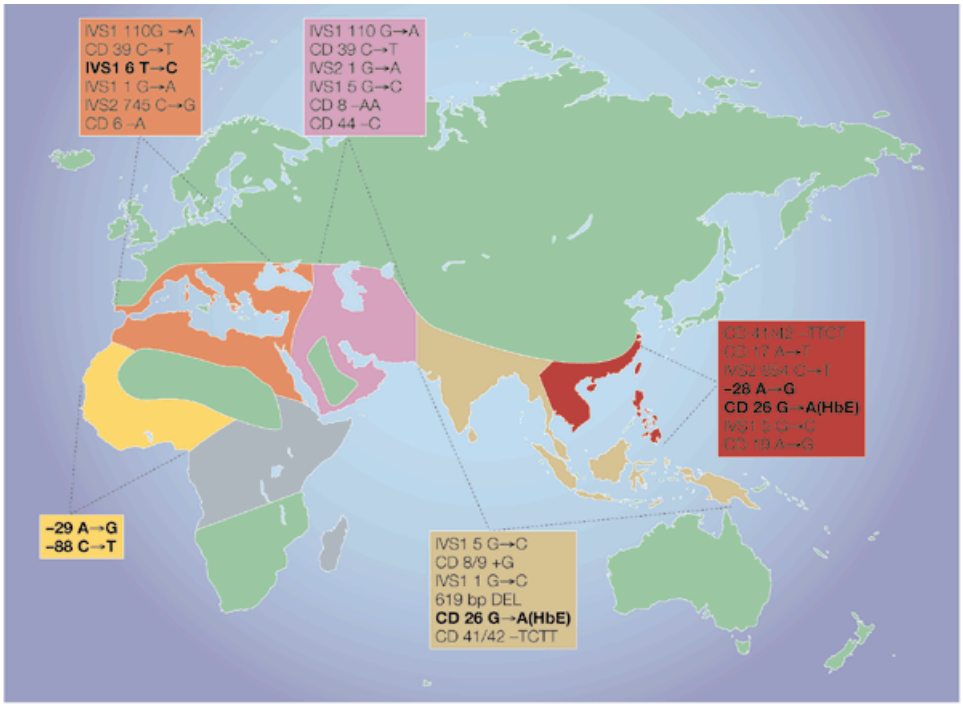


Figure 1.4 - Global β-thalassaemia distribution. The common mild mutations are shown in bold. β-thalassaemia also occurs in the regions shaded in grey, but little is known about its molecular pathology in these areas. Mutations vary according to different geographic zones, mutation variation is shown in different colours (Image taken from Weatherall, 2001).

1.1.5 What does the future hold for thalassaemia patients?

Recent medical advances bought a new hope for thalassaemia patients with iron chelation and bone marrow transplantation therapy. However, treatment options are still far from satisfactory; as blood transfusions entail frequent disturbance for the patient with regular hospital visits and allogenic BMT is still far for from being a feasible option for most patients. Furthermore, iron overload leading to cardiac disease remains the main cause of death in thalassaemia patients.

Advances in genetic therapeutics open a new door for thalassaemia patients, offering the hope of new, alternative treatment requiring less hospital dependence and life-long disease correction.

1.2 Regulation of *HBB* expression

The development of gene therapy strategies for the haemoglobinopathies requires a clear understanding of the molecular mechanisms of gene expression of the *HBB* locus.

The *HBB*-like genes are arranged in a linear array on the short arm of chromosome 11. The cluster consists of five active genes (*HBE*, *HBG-G*, *HBG-A*, *HBD*, and *HBB*) arranged in the order that reflects their expression during development (**Figure 1.5**, left hand panel) (Sadelain, 2006).

The embryonic ε-globin gene (*HBE*) is transcribed in the embryonic yolk sac in the first (6-8) weeks of gestation and is located at the 5' end of the cluster. After the switch in the site of haematopoiesis from the yolk sack to the foetal liver, *HBE* is repressed and the two γ-globin genes (*HBGy-Ay*) located downstream of *HBE* are expressed. These two genes are expressed in very low quantities in the yolk-sack and mainly in the foetal liver derived erythrocytes until birth. In the second switch, which occurs shortly before birth and lasts until BM becomes the major site of haematopoiesis, coincides with the activation of the adult *HBB* gene at the 3' end of the cluster. *HBG* is also activated in erythroid cells derived from BM haematopoiesis although expressed at a much lower level (approximately 1% of total haemoglobin) (Levings & Bungert *2002).*

Erythroid specific promoters and enhancer elements have been described for each of the active genes. Transcriptional regulation that leads to differentiation and developmental stage specific expression in the *HBB* locus is mediated by DNA-regulatory sequences located both proximal and distal to the gene cluster (Antoniou & Grosveld, 1999).

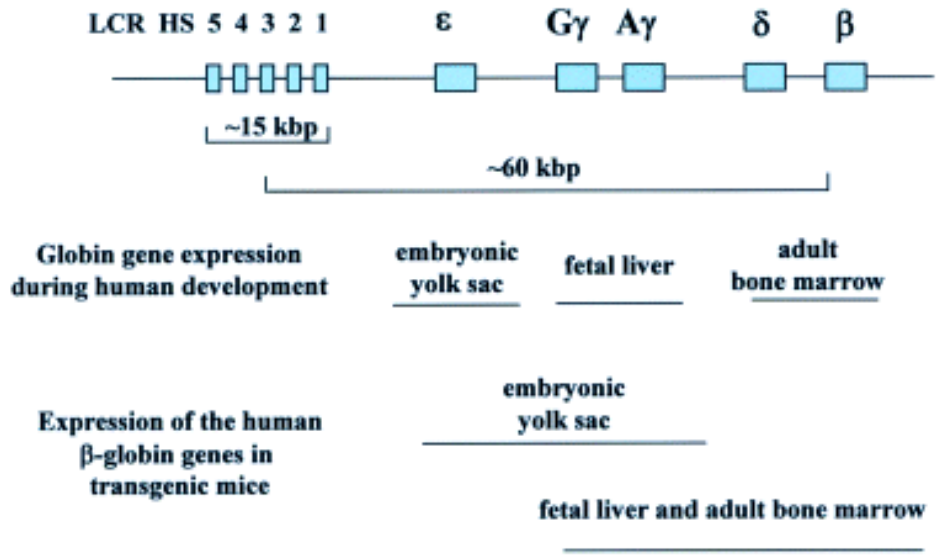


Figure 1.5 - The structure of human *HBB* locus. The five genes of the human *HBB*-like gene family are arranged in linear order reflecting their expression during development. The image also shows the DNaseI hypersensitive site core elements that constitute the locus control region (LCR) regulatory element comprising 200-400bp. (Image taken from Levings & Bunget, 2002)

1.2.1 Introduction to the *HBB* locus control region

The most potent distal regulatory element in the *HBB* locus is the locus control region (βLCR), located within a 16kb stretch of DNA upstream of the *HBB*-like genes. The βLCR functions both as a powerful tissue specific enhancer and dominant chromatin remodelling regulatory element, able to overcome chromatin position effects and confer upon a gene linked in *cis*, physiological levels of expression directly proportional to transgene copy in mice (Fraser & Grosveld 1998, Li *et al.*, 1999). This has been shown to be the major *cis* regulator of all five *HBB*-like genes, since its absence results in transcriptional suppression of all genes at all stages of development (Feng *et al.*, 2005).

The βLCR is comprised of five separate 200-400bp core sequences that exhibit both erythroid-specific and ubiquitous transcription factor binding sequences together with extreme sensitivity to DNase I digestion in erythroid cells (Bulger &

Groudine, 1999). These were identified as DNaseI hypersensitive sites (HS) in nuclear chromatin. In fact, the entire *HBB* locus remains in an inactive DNaseI - resistant chromatin conformation in cells within which these globin genes are not expressed. Studies have shown that sensitivity to DNaseI within the *HBB* locus in erythroid cells varies and depends upon the developmental stage of haematopoiesis (yolk sac, foetal liver, adult spleen)

It has been found that deletions of the βLCR will lead to repression of *HBB* expression (Antoniou & Grosveld, 1999).Therefore, the βLCR acts as an essential *cis*-acting regulatory element that is required for expression of all genes from the *HBB* locus *in vivo*.

The overall organisation and sequence of the βLCR is conserved among several vertebrate species, adding further support to the vitally important role that this element plays in the regulation of expression of the *HBB*-like gene cluster (Hardison *et al.,* 1997).

1.2.2 Properties of the βLCR

The mechanism of activation and silencing of native *HBB* genes and transgenes during development has been under intense investigation for many years.

The *HBB* promoter strongly depends on the βLCR for high expression as its transcriptional activity is massively reduced (<1%) if the βLCR is deleted. In fact, globin transgene constructs reveal strong position-of-integration effects in transgenic assays in the absence of the βLCR. The incorporation of the βLCR into these constructs has been shown to confer a chromatin remodelling capability and enhances globin gene expression maintaining a transcriptionally competent domain within an erythroid environment. LCRs are associated with tissue-specifically expressed genes and have a strong transcriptional enhancer regulatory function and also a dominant chromatin remodelling capability (Antoniou & Grosveld, 1999).

The enhancer activity of the βLCR is erythroid specific and studies have shown that this resides mainly in 5' HS2. Specifically, HS2 behaves as a classical enhancer and its activity can be detected in transient transfection assays, whereas transcriptional enhancing activity from HS3 and HS4 can only be detected after integration, within chromatin. It is generally considered that HS3 acts as the core chromatin opening element, as this is the only element that functions as a single copy transgene in mice, while HS1, 2 and 4 act as LCR-type elements only when present as multiple copies. This suggests that the βLCR consists of two types of elements; the "core", with the capacity of chromatin opening and the other HS sites - powerful enhancer elements that boost the transcriptional potency of HS3 (Antoniou & Grosveld, 1999). HS5 functions as a stage-specific (embryonic) insulator element (Li *et al.*, 2002)

The enhancer function of the βLCR is thought to be mainly mediated by transcription factors binding to the core regions of the DNase I HS sites 1-5. HS1 to 4 contain an array of ubiquitous and erythroid specific transfactor binding sites, with a conserved region within HS2 (TGCTGA(C/G)TCA(T/C) thought to be critical for strong enhancer activity This region is known to be a Maf recognition element (MARE) bound by multiple homodimeric or heterodimeric transcription factors *in vivo* (Talbot & Grosveld, 1991; Li *et al.*, 1999).

The βLCR is not the key factor in tissue specificity, which depends on both the LCR and globin promoters (Tolhuis *et al.*, 2002; Li *et al.*, 2002; Carter *et al.* 2002). In addition, the βLCR is also able to enhance the expression of linked heterologous non-globin gene promoters in erythroid cells. It can therefore be concluded that tissue specific control of basal transcription resides in the *HBB* promoter, whereas the tissue specific enhancement of gene expression depends on the βLCR.

1.2.3 Mechanisms of *HBB* activation

There have been four possible models of LCR-mediated gene expression proposed; looping, tracking, facilitated tracking and linking (Li *et al.*, 2002). Several lines of direct experimental evidence have now shown the looping model to be correct (Palstra *et al*., 2008).

The looping mode suggests that the 5'HSs of the β-globin LCR fold to form a holocomplex (Carter *et al.* 2002 Tolhuis *et al.,* 2002).

Initial studies performed using RNA-FISH TRAP provided evidence that the βLCR HS1-HS4 exist in physical proximity to the actively transcribed *Hbb* genes with the intervening DNA looping out in mouse erythroid cells, with the HS2 site the one in most intimate interaction (Carter *et al.* 2002). Similarly, chromosome conformation capture (C3) performed by *Tolhuis* and colleagues showed that the murine *Hbb* LCR is in physical proximity to the active globin gene *in vivo* in expressing tissue with the intervening DNA looping out. Experiments revealed that there are cell-specific interactions within the *Hbb* locus and that the DNase I HS sites at both ends of the locus participate in these interactions. Interactions and looping were not observed in non-expressing tissue (brain) which led to the conclusion that multiple sites acting over 130kb interact to form a cluster in the nuclear space and this is crucial in establishing an open chromatin domain and activating transcription, a structure coined as the "active chromatin hub" (ACH) (Tolhuis *et al.*, 2002).

Together, these studies further support the looping model in that they show the βLCR is only able to interact with one promoter at a time but it is free to move from gene to gene in a "flip-flop" mechanism between two or more promoters depending on the developmental stage, causing stable enhancer-promoter interactions - a key structure that sets up autonomous expression of genes irrespective of their position in the genome (Levings & Bungert 2002). The chromatin looping mechanism is represented in **Figure 1.6**.

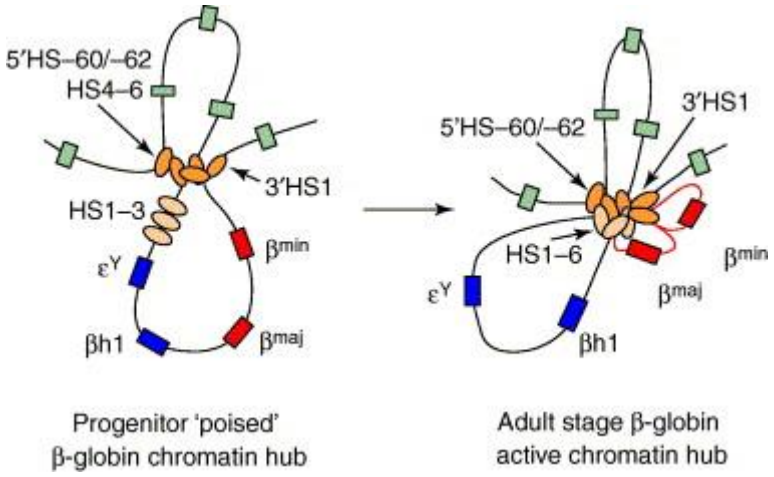


Figure 1.6 - Chromatin looping in the mouse *Hbb* locus. In erythroid progenitor (left) a chromatin hub forms consisting of HS1 and HS4. In differentiated erythroid cells (right) an active chromatin hub incorporates the remaining LCR HS1-3. The globin genes join the holocomplex when they are being actively transcribed leaving the inactive genes looped away (Image taken from Dean, 2006).

.

The formation of a chromatin hub (CH) initially occurs through the interaction between CCCTC-binding factor (CTCF)-bound *cis*-regulatory elements, a transcription factor involved in gene insulation and boundary formation at transcription competent domain. Subsequently, the remaining HSs of LCR (HS1-3) interact with the CH to form an ACH to activate transcription of the β-globin gene family members stably at different points during development (Palstra *et al.*, 2008).

Close association with nearby promoter and enhancer elements allows the delivery of LCR-bound transcription proteins and coactivators hat interact with the basal transcription apparatus to form a stable transcription complex, thus enhancing globin gene expression (Li *et al.*, 2002).

Erythroid specific transcription factors, such as ELKF and GATA-1 are involved in ACH stabilisation and LCR-promoter interaction respectively (Dean, 2006; Palstra *et al.*, 2008).

1.2.4 Incorporation of the *HBB* elements into delivery vectors

Despite their chromatin remodelling capabilities, the individual βLCR elements used on their own do not completely negate silencing of a transgene and the use of multiple HS sites is required for maximum activity, which causes problems in viral vectors of limited cloning capacity. The use of severely paired down core regions can be detrimental to their function, resulting in only partial negation of insertion-site position effects (Sadelain, 2006).

Furthermore, it is known that the ACH model of enhancer-mediated promoter activation requires a minimum distance between enhancer and gene to allow DNA looping for full transcriptional activation (de Laat & Grosveld, 2003). This mechanism is therefore severely compromised in a compact, size constrained viral vector.

The discoveries of globin controlling elements and the recent advances in globin regulation have contributed enormously for the development of efficient gene therapy vectors regulating *HBB* expression (see below). These functional observations represent key pieces in the erythroid-specific gene therapy puzzle, of which we are finally starting to perceive a clear image.

1.2.5 HBB splicing mechanisms and the *HBB* expression cassette design

The protein coding regions of most eukaryotic genes are interrupted by intervening sequences or introns. These intronic sequences are subsequently removed or spliced from the primary transcript as part of the maturation process to mRNA. Interestingly, however, the removal of introns from genes and their subsequent expression in both tissue culture cells and transgenic mice, results in very low levels of mRNA, which is not due to transcriptional insufficiency resulting from the absence of an intronic enhancer element.

HBB has a three exon-two intron structure (**Figure 1.7**) and gene expression and mRNA levels decrease after splicing. Physiological levels of expression of *HBB* from a natural chromatin context by stable transfection in murine erythroleukaemia (MEL) cells under the control of the βLCR, can only be achieved in the presence of the second intron (Antoniou *et al.*, 1998).

Studies suggest that the second intron of *HBB* is essential in promoting cleavage and polyadenylation (Antoniou *et al.*, 1998; Millevoi *et al.*, 2002). The efficiency of normal 3'-end formation remained the same even when intron 2 is placed in the intron 1 position. But the efficiency is reduced to 16% and 30% when the second intron is replaced by other heterologous introns namely intron 1 or α-globin intron 2, respectively (Antoniou *et al.*, 1998).

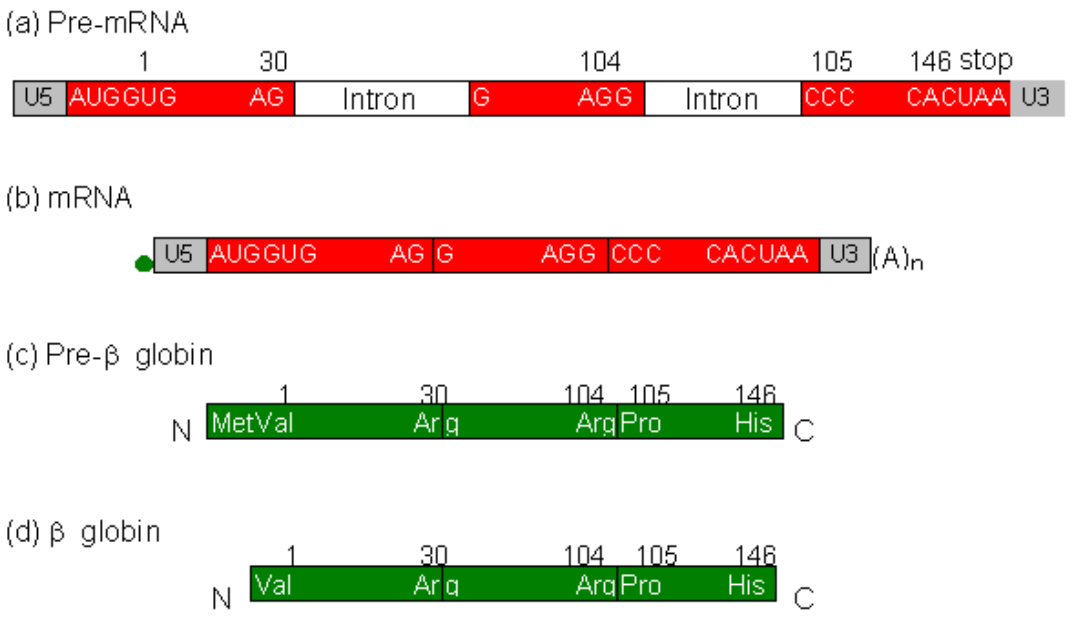


Figure 1.7 - Expression of *HBB* gene. U5 and U3 represent untranslated regions at the 5' and 3' end, respectively.
(Image taken from http://www.web-books.com/MoBio/Free/Ch5A4.htm).

1.2.5.1 Additional elements to increase expression from the *HBB* cassette: the *HBB* transcription termination region (βterm)

Transcription termination in eukaryotes involves cleavage of the new transcript followed by template-independent addition of adenosine residues at its new 3' end, in a process called polyadenylation. Termination is triggered following recognition of the poly(A)-addition (pA) signal by Pol II and subsequent pre-mRNA cleavage, which occurs either at the pA site or in transcripts at terminator regions. Failure in Pol II to terminate efficiently leads to inefficient pre-mRNA cleavage and polyadenylation, resulting in low gene expression (West & Proudfoot, 2009).

Recent studies reported a transcriptional termination region located 600-1200bp 3' of the pA site of *HBB*, capable of increasing the efficiency of polyadenylation and pre-mRNA cleavage (West *et al.*, 2008; Dye & Proudfoot, 2001) – the HBB transcription termination region (βterm). Studies have shown that *HBB* transcription termination releases transcripts from the DNA template, which subsequently results in their processing and accumulation of stable cytoplasmic mRNA (Dye & Proudfoot, 2001). The authors later reported a 10-fold HBB protein increase in HeLa cell pools in the presence of βterm (West & Proudfoot, 2009).

During *HBB* transcription termination, the pA signal switches Pol II configuration after passing through the pA site, where the Pol II-associated unprocessed pre-mRNA is cleaved and released from the template. Finally, the cleavage occurs at the poly(A)-addition site and dissociates pre-mRNA from Pol II.

Low efficiency of polyadenylation results into high amounts of unspliced pre-mRNA accumulation and degradation in the cell nucleus. Transcription termination is thus a determinant of the efficiency of pre-mRNA 3′ end processing and the βterm region is a key element for efficient HBB pre-mRNA cleavage, polyadenylation and gene expression (**Figure 1.8**; West & Proudfoot, 2009).

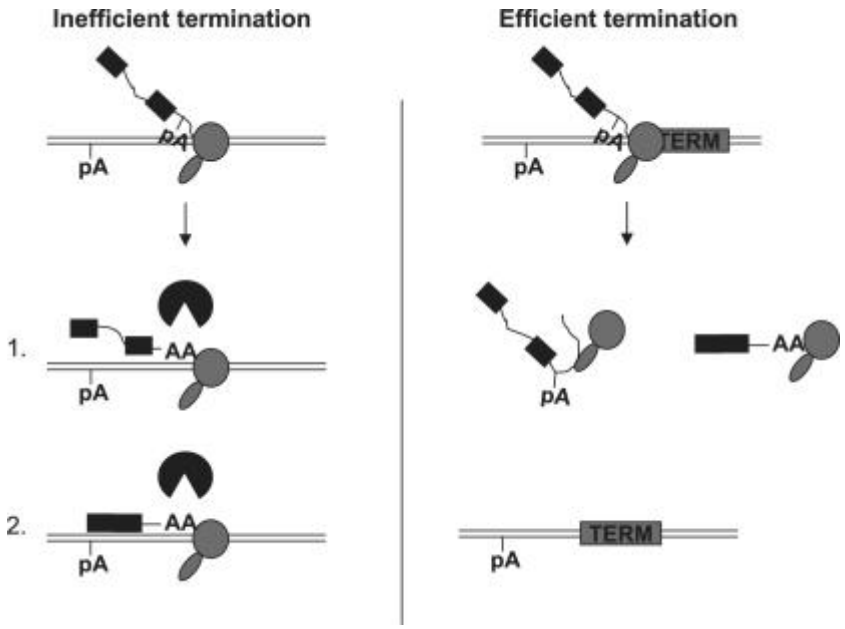


Figure 1.8 – Role of the *HBB* transcription termination region (βterm) in pre-mRNA cleavage and polyadenylation. Left panel; when termination is inefficient, pre-mRNA splicing is compromised and polyadenylated RNA species (1 and 2) are retained at transcription sites. These events are surveyed by the nuclear exosome, which degrades aberrant RNA and reduces gene expression. Right panel ; termination enhances gene expression by promoting more efficient pre-mRNA processing and so reducing the susceptibility of transcripts to degradation. (taken from West & Proudfoot, 2009)

1.3 Gene therapy

Gene therapy can be broadly defined as the treatment of a disease or medical disorder by the introduction of therapeutic genes into the appropriate cellular targets. These therapeutic genes can correct deleterious consequences of specific gene mutations or re-program cell functions to overcome a disease. For successful gene therapy, the exogenous therapeutic gene must be specific, efficient, and stably incorporated into the target cell.

Gene therapy can actually be performed by both “viral delivery”, taking advantage of the infectious nature of viruses and “non-viral” by the use of naked DNA pharmaceutical formulations (for example, liposomes) into a cell.

1.3.1 Non-viral gene therapy

Non-viral methods are essentially focused on the incorporation of “naked DNA” and oligonucleotide structures into the cell. This methodology presents certain advantages over viral methods, such as simple large scale production and low host immunogenicity. However, to date, non-viral gene therapy has failed to provide sufficient evidence of stable expression and delivery to be competitive over viral delivery. This report will focus on viral gene therapy and the use of lentiviral vectors for gene delivery.

1.3.2 Viral gene therapy

Viruses are intracellular obligate parasites and they have evolved as efficient vehicles for the delivery of DNA or RNA to target cells. Wild-type viruses infect their hosts and introduce their genetic material into the host cell as part of their replication cycle. This genetic material, hijacks the host cell normal replication machinery to serve the needs of the virus and the host cells eventually follow these instructions and produce additional copies of virus, leading to more and more cells becoming infected. This infection pattern leaded to the hypothesis that viruses could be used as vehicles

to carry 'good' genes into a human cell. The removal of viral genes that induce the disease from the virus and the replacement of these with genes encoding the desired therapeutic product proved to be a viable and safe alternative to common therapeutics, and the use of viral mediated delivery opened a new door in the gene therapy field.

To date, a large number of viruses with unique characteristics useful for gene therapy have been identified. This has led to the application of recombinant viruses such as adenoviruses, adeno-associated viruses, herpes viruses, poxviruses, retroviruses, and, more recently, lentiviruses, both in the laboratory and clinic.

1.3.2.1 Adenovirus

Adenoviruses are non-enveloped viruses containing a linear double stranded DNA genome. To date over 40 serotype strains of adenovirus have been described, most of which cause benign respiratory tract infections in humans. The adenovirus life cycle does not normally involve integration into the host genome, instead they replicate as episomal elements in the nucleus of the host cell. Adenovirus wild type genome is approximately 35 kilobases of which up to 30 kb can be replaced with foreign DNA (Smith, 1995; Verma & Somia, 1997). This large vector capacity is being fully exploited by gene therapists.

Adenoviral vectors are very efficient at transducing target cells *in vitro* and *in vivo*, and can be produced at high titres ($>10^{11}$/ml), However, transgene expression *in vivo* from progenitor vectors tends to be transient, limiting therapeutic applications of the vector (Verma & Somia, 1997).

1.3.2.2 Adeno-associated virus

Adeno-associated viruses (AAV) are non-pathogenic human parvoviruses dependant on a helper virus, usually adenovirus, to proliferate. Their wild type genome is a single stranded DNA molecule, consisting of two genes; *rep*, coding for proteins which control viral replication, structural gene expression and integration into the host genome; and *cap*, which codes for capsid structural proteins. These viruses infect both dividing and non-dividing cells, and in the absence of a helper

virus mostly integrate into a specific point of the host genome (19q 13-qter) at a high frequency (Kotin *et al*., 1990) proving stable and controlled expression.

AAV gene therapy vectors can infect both dividing and quiescent cells and persist in an extrachromosomal state without integrating into the genome of the host cell. Success with these vectors has been reported recently for the correction of the retina(Bainbridge *et al.,* 2006, Alexander *et al.* 2007).

1.3.2.3 Retrovirus

Retroviridae is a family of single-stranded (ss) RNA spherical viruses of around 80-120 nm in diameter (Vogt & Simon, 1999). This family of viruses can be distinguished by three main characteristics: genetic information in RNA form; virions that possess reverse transcriptase enzyme; and a virion morphology that consists of two proteinaceous structures and a dense core, all surrounded by an envelope structure.

The wild-type retrovirus consists of a lipid bilayer-enveloped particle comprising a homodimer of linear, positive sense, ssRNA genomes between 7 and 11 kb in length. Following entry into target cells, retroviruses perform a journey from the cell surface to the nucleus while the RNA genome is retro-transcribed into linear double stranded (ds) cDNA (termed provirus) via reverse transcriptase enzyme. Once inside the nucleus, the provirus DNA integrates into the host cell genome via the action of *integrase* enzyme and is passed on to daughter cells upon cell division.

The ssRNA genomes are converted to dsDNA and integrate, via long terminal repeats (LTR) located at both ends (3’ and 5’) in the target cell genome. The LTRs represent the two end parts of the viral genome that are connected to the cellular DNA of the host cell after integration consisting of 3’ unique elements (U3), repeat elements (R) and 5’ unique elements (U5). These do not encode for viral proteins but harbour cis-acting elements important for provirus integration. Viral LTRs frame the coding genes; the tandem *group-specific antigen* (*gag*) gene coding for core and structural proteins of the matrix and nucleocapsid core, the *polymerase* (*pol*) gene coding for enzyme reverse transcriptase protease and integrase; and the *envelope* (*env*) gene coding for the retroviral coat proteins and containing information for the

surface and transmembrane components of the envelope protein. The "classical" structural scheme of a retroviral genome is: 5'LTR-gag-pol-env-LTR 3 as illustrated of **Figure 1.9**. Other important components are the packaging signal (psi or ψ), required for the specific packaging of the viral RNA genome into newly formed virions (Watanabe & Temin, 1982), and the central polypurine tract (cPPT), which is the site of the initiation of positive-strand DNA synthesis during reverse transcription (Charneau *et al.*, 1992; Rattray & Champoux, 1989).

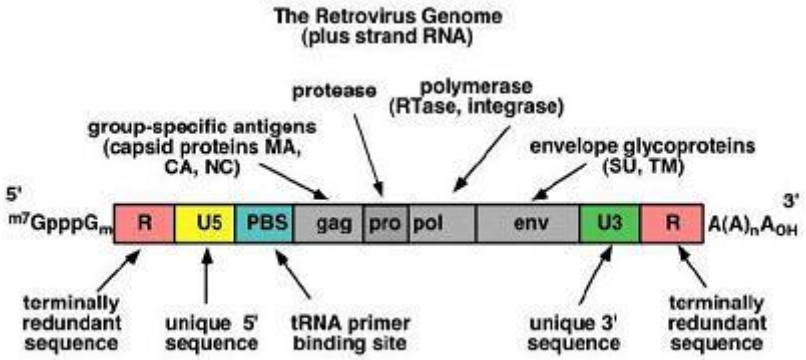


Figure 1.9 - General structure of an integrated retroviral genome. The long terminal repeats (LTRs) have sequences needed for the regulation and initiation of transcription. The sequence denoted ψ is required for packaging retroviral RNAs into mature virus particles. (Image taken from *http://envmedical.com*)

Based on their genome organization, the *Retroviridae* are divided in simple and complex retroviruses, examples of which are oncoretroviruses, such as MLV, and lentiviruses, such as human immunodeficiency virus (HIV)-1, respectively. Lentiviruses include primate (human immunodeficiency virus HIV and simian immunodeficiency virus SIV) and non-primate (feline immunodeficiency virus FIV, bovine immunodeficiency virus BIV, caprine arthritis-encephalitis virus CAEV, equine infectious anaemia virus EIAV and visnavirus) retroviruses.

1.3.3 Lentivirus

Lentivirus infection is characterized by a long interval between the initial infection and the onset of serious symptoms, hence its name comes from the latin terms *lent, latent* meaning "slow" viruses. Lentiviruses are a subfamily and have more complex genomes than oncoretroviruses. This report will focus on human immunodeficiency virus (HIV).

1.3.3.1 HIV virus

1.3.3.1.1 HIV genome structure

The HIV wild type genome consists of two identical unspliced ssRNA molecules of approximately 10 kb long located within a nuclear core surrounded by a nucleocapsid and a viral envelope core.

Their genome is organized into the *gag, pol,* and *env* gene configuration (**Figure 1.10**). Gag encodes the structural proteins, whereas the pol gene encodes the enzymes that accompany the ssRNA. Of these, reverse transcriptase carries out reverse transcription of the viral RNA to DNA, integrase catalyses the integration of the proviral DNA into the host genome, and protease is involved in *gag/pol* cleavage and virion maturation (Katz & Skalka, 1994). Env encodes the viral envelope.

Wild type HIV-1 contains two regulatory genes, *tat* and *rev*, essential for viral replication and gene expression. Tat protein activates the promoter in the HIV LTR so that viral RNA is produced efficiently, and Rev interacts with a region of viral RNA known as the Rev-responsive element (*rre*) and promotes the transport of viral RNA from the nucleus to the cytoplasm (Pluta & Kacprzak, 2009)

Although possessing the retroviral basic genome organisation, HIV-1 specifically contains four additional accessory genes, *vif, vpr, vpu* and *nef,* critical for *in vivo* replication and pathogenesis (**Figure 1.10**; Coil & Miller, 2004). The presence of the latter in therapeutic vectors raises safety concerns as the proteins they encode have cytotoxic or cytostatic activities; *vpr* induces G2 cell cycle arrest and *nef* alters

cellular activation pathways. Cell surface molecules such as CD4 and the class I major histocompatibility complex are down-regulated by *nef* and *vpu*. Moreover, *nef*, *vif* and *vpr* are incorporated into viral particles and can enhance the immunogenicity of vectors. It is therefore important to ensure that lentiviral vectors (LV) are constructed in the absence of these genes, which have been shown not to be essential for viral transduction (Delenda, 2004; Pluta & Kacprzak, 2009).

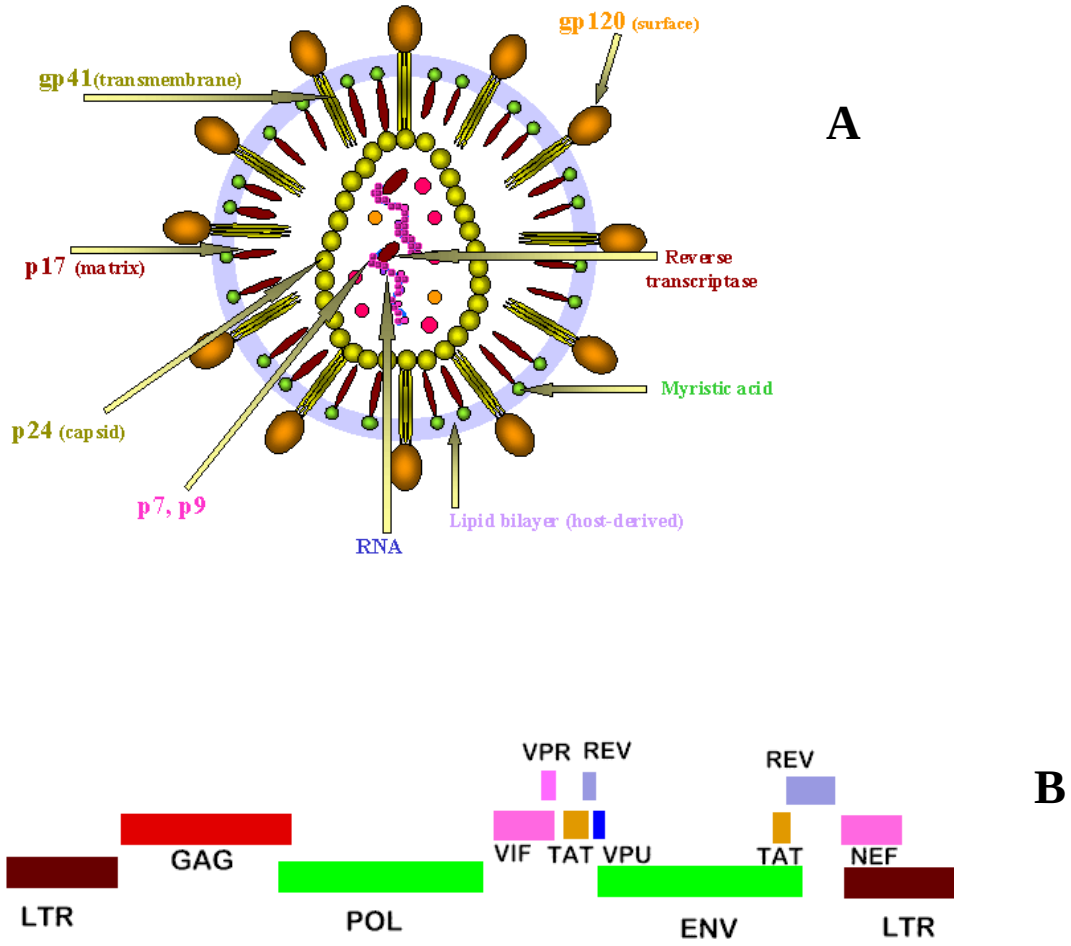


Figure 1.10 – The HIV-1 wild type genome. A) shows the spatial representation of the virus, with the encoding proteins (gp41, p17, p24, p7.,etc.) surrounding te core containing the viral RNA and reverse transcriptase. The virus is mainly composed by three genes (*Gag*, *Pol* and *Env*) encoding the packaging and envelope proteins of the virus flanked by the two LTRs, as shown in **B)** (Image from *http://pathmicro.med.sc.edu/lecture/hiv9.htm*)

1.3.3.1.2 HIV life cycle

HIV replication starts when the viral envelope glycoprotein gp120 binds to a CD4 receptor and a secondary receptor on the cell surface after which conformational changes in the non-covalently associated gp41 subunit releases free energy sufficient to promote fusion of the viral particle envelope with the cell membrane. The viral core is then spilled into the cytoplasm and reverse transcription of the viral RNA occurs. The reverse transcription complex (RTC) is then transported towards the nucleus via the cellular protein dynein. When reverse transcription is complete and the provirus and several viral proteins form, the pre-integration complex (PIC) is allowed entry into the nucleus via the cellular protein LEDGE. After integration, transcription is conducted by RNAP II and enhanced by the viral protein Tat. Unspliced and partially spliced transcripts are stabilised by the viral protein *Rev* (Pluta & Kacprzak, 2009).

1.3.3.2 Development of lentiviral vectors

To minimize the risk associated with manufacturing and use of LVs, all non-essential genes coding for accessory proteins and responsible for virulence, should be removed from the vector sequence. Additionally the vector genome is usually split into several parts with limited sequence overlap to reduce to a minimum the possibility of independent recombination to form wild type viral particles and vector mobilization. The genes required for lentiviral vector generation are therefore separated on different plasmids that are co-transfected into HEK293T cells together as a complex aggregated with the help of a transfection agent (see *Materials and Methods section*). Hence, such a system typically consists of packaging expression cassettes (helper), a vector cassette (transfer vector) and an envelope expression cassette (**Figure 1.11**).

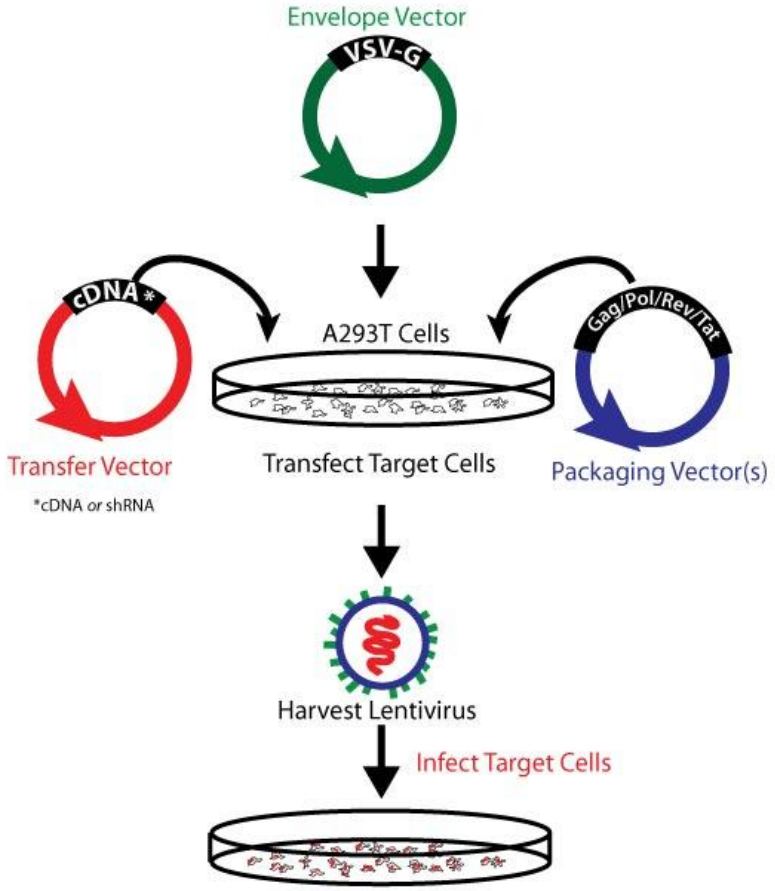


Figure 1.11 - Lentiviral vector production protocol. As in oncoretroviral vectors, the viral genes *gag*, *pol* and *env* are replaced by the transgene of interest. The LTR and the packaging signal ψ are maintained for vector amplification and packaging. Of the lentiviral accessory proteins, only the *rev* responsive element (RRE) is maintained as an additional *cis*-acting sequence required for the nuclear export of unspliced and single-spliced viral RNAs in the presence of the Rev protein. Rev is expressed by one of the three packaging constructs. The two others encode the Gag/Pol or the Env proteins, respectively.
(Image taken from http://www.addgene.org/)

The development of a stable cell line producing LVs has been reported. These STAR cells are derived from 293T or HT1080 human cell lines that were transduced with an MLV-based vector expressing a codon optimised version of *gag* and *pol* (lacking RRE) and then co-transduced with Tat and Rev being expressed from separate gammaretroviral vectors. The STAR cells were then stably transfected with plasmids expressing either MLV-A, GALV+, or RD114pro glycoproteins to allow for high plasticity regarding vector pseudotyping. The production of SIN and non-SIN versions of LVs in these cells is similar, with a high titre and stable for up to 3 months (Ikeda *et al*., 2003).

1.3.3.2.1 Transfer plasmid

The transfer vector plasmid contains the full-length vector RNA containing *cis*-acting elements for efficient packaging, reverse transcription, nuclear import and integration into the host genome. Typically, it contains the transgene expression cassette flanked with the LTRs at both ends of the genome, as in wild type virus, the neighbouring sequences act as *cis*-elements that regulate the virus during viral gene expression, packaging, transcription and integration, being essential for the viral infection of the host cell and genome integration.

The expression cassette contains a gene of interest driven by an internal promoter usually positioned between the 3' Tat/Rev splice acceptor site and the 5' LTR and contain non-coding *cis*-acting regulatory sequences needed for packaging and to increase the efficiency of gene transfer. The promoter can be constitutively active and derived from a virus [such as the LTR from the spleen focus-forming virus (SFFV), or the immediate early promoter from the cytomegalovirus (CMV)] or cellular genes [such as the phosphoglycerate kinase (PGK) or elongation factor 1α (EF1 α)]. These promoters are active in a range of different cell types. However, transcription can be restricted to a given cell type by the use of tissue-specific promoters, such as the human desmin promoter in skeletal muscle (Talbot *et al*., 2010). Alternatively, the incorporation of microRNA target sequences in the transcript can also prevent expression in cells where those microRNAs are expressed (Brown *et al*., 2006). Multiple promoters can be included, each driving expression of

a different transgene. However, this strategy has the potential problem of promoter suppression, where one of the transcripts is expressed more than the other (Emerman & Temin, 1984).

The packaging signal (ψ or E) is located near the 5' end of the genome between the splice donor and the gag start codon, the primer binding site (PBS), important on first strand synthesis, and the 3' and 5' LTRs.

Additional modifications to this transfer vector included the incorporation of the central polypurine tract (cPPT) to enable the formation of a DNA flap during reverse transcription, facilitating nuclear translocation of the PIC in both dividing and non-dividing cells and providing a second site for initiation of DNA synthesis. This DNA flap has been shown to increase nuclear import of the viral DNA, thus increasing the transduction efficiency in both dividing and non-dividing cells (Follenzi *et al*., 2000; Zennou *et al*., 2000).

1.3.3.2.2 Packaging plasmid

The packaging cassette expresses viral enzymes and structural proteins necessary for infectious particle formation.

The original first generation LV system developed in 1996 contained the accessory proteins Vpu, Vpr, Vif Nef and the Rev and Tat regulatory proteins and functional components of the packaging plasmid. The evolution of the packaging cassette design and optimisation of the transfer vector eventually led to the removal of five proteins associated with HIV virulence - tat, nef, vif, vpr, and vpu.

These packaging constructs were further refined by eliminating all accessory proteins that are associated with virulence and toxicity and are not required for viral replication, eventually leading to the development of the second-generation packaging cassette (Zufferey *et al*., 1997). In this system only Rev and Tat proteins were expressed together with gag and gag-pol polyproteins.

The packaging plasmid contains also the *gag* and *pol* genes, driven by a heterologous promoter. In order to improve the biosafety of the system and overcome the problem of susceptible points for homologous viral recombination between the packaging and the transfer vector constructs, *rev* was later placed on a second plasmid while *tat* was completely removed and its function replaced using modified 5' LTR elements containing strong constitutive RSV or CMV-derived promoters in the corresponding vector constructs - these systems are often referred to as a third generation packaging cassette.

Safety improvement in the design of the lentiviral packaging system is still an area of intensive research. Attempts to create hybrid LV systems exploiting the limited sequence homology between other viruses have been pursued. It was shown tat HIV-1 vectors can be packaged with SIV core particles and permutations between HIV-1 and HIV-2 packaging have been tested. Furthermore, lentiviral particles containing packaging genes from a recombinant adenoviral vector have been designed (Kuate *et al.*, 2004).

The packaging plasmid contains the integrase protein responsible and essential for LV integration into the host cell genome. Mutations performed on this protein have given rise to the non-integrating packaging plasmids (see *Section 1.3.2.8*).

1.3.3.2.3 Envelope plasmid

The HIV wild type *env* shows preferable transduction of $CD4^+$ cells. As this protein is removed from the system, an additional plasmid expressing a heterologous glycoprotein is used during vector production. The envelope plasmid contains the gene involved in target cell recognition, the *env* gene. The normal HIV envelope is usually exchanged with glycoproteins from other viruses so that the vector tropism is extended to a wider range of cells other than lymphocytes. Pseudotyping of this type offers several advantages: increase of vector safety due to the elimination of sequence homology with the wild type virus, expands tropism of the pseudotyped vector with

the targets cells, and improves vector stability allowing the production and concentration by centrifugation of high titre vectors of LV to produce the robust nature of the envelope and long-term storage.

The most commonly used glycoprotein to pseudotype LVs in order to expand vector tropism is the glycoprotein from the vesicular stomatitis virus (VSVg). The receptor for VSVg, although still undetermined, appears to be ubiquitous in all cell types, explaining the broad host range of VSV (Schlegel *et al*., 1982; Coil and Miller, 2004). Furthermore, VSVg pseudotyped vectors are highly stable and can be efficiently concentrated by ultracentrifugation producing unequalled titres of at least 10^7-10^8 transducing units (TU) per ml, and enabling the production of serum-free, high-titre vector particles (Burns *et al*., 1993, Naldini 1996, Reiser *et al.,* 1996, Poeschla *et al.,* 1996). This envelope facilitates vector entry via the endocytic pathway which diminishes the requirement for viral accessory proteins for full infectivity. However, the mechanism responsible for cell binding and well as the cellular receptors involved in the glycoprotein cell recognition remain unclear.

One drawback of the production of pseudotyped vectors using VSVg is that due to its fusogenic properties it is toxic to some mammalian cells if constitutively expressed (Burns *et al.,* 1993). This means that producer cell lines expressing packaging proteins for long-term virus production require conditional production of this glycoprotein. Another drawback of this pseudotype is inactivation by human serum complement, preventing its use *in vivo* (DePolo *et al*., 2000). Furthermore, it has been reported that LV preparations pseudotyped with VSVg are contaminated with tubulovesicular structures carrying nucleic acids or proteins that may elicit an undesirable immune response (Pichlmair *et al*., 2007).

Other than VSVg, otherheterologous glycoproteins have also been used to restrict the tropism of LVs. The neurotropic properties of Lyssavirus, including the rabies and Mokola virus, have been studied. Vectors pseudotyped with those glycoproteins infect neurons preferentially (Mochizuki *et al*., 1998; Desmaris *et al*., 2001). Retrograde transport has been achieved by the use of Rabies glycoprotein (Mazarakis *et al*., 2001).

Another strategy to target specific cell types is the engineering of envelope glycoproteins. Viral envelope proteins have been modified to contain ligands or single chain antibodies (Hatziioannou *et al*., 1999). Using the second approach, by fusing single chain antibodies to the Sindbis glycoprotein, LVs have efficiently been targeted to melanoma cells (Morizono *et al*., 2005).

1.3.3.3 Development of SIN vectors

Virologists are well aware of the risks associated with viral recombination that can theoretically occur at the DNA level, especially during the mixing of co-transfected plasmids, between transfected plasmids or between the proviral vector and a homologous chromosomal sequence in the target cell. A major issue regarding the biosafety of LVs remained: the possibility of transfer vector retaining the ability for transcription of the full-length genome after integration into the target cell.

It is possible for the LV provirus to be mobilised by replication competent virus (for example, if the transduced cell were subsequently infected with wild type HIV-1). In addition, there is the possibility that the 3' LTR could induce aberrant expression of adjacent genes.

The first self-inactivating (SIN) retroviral vectors were made in an MLV-based vector (Yu *et al*., 1986; Yee *et al*., 1987). The first vector had a deletion of 299bp in the U3 region, containing the enhancer and promoter CAAT box element with the second having the additional deletion of the TATA box. However, retroviral vectors display weak polyadenylation sites and deletion of U3 to generate SIN vectors increased the likelihood of read-through thus augmenting the potential for insertional mutagenesis (Furger *et al*., 2001; Zaiss *et al*., 2002).

The same principle was used later in LVs design by *Miyoshi* and colleagues *(*Miyoshi *et al.,* 1998) by deletion of a 133bp section from the U3 region of the 3' LTR, which removed the TATA box and other transcription factor binding sites, resulting in transcriptional inactivation of the proviral LTR in infected cells, both *in vitro* and *in vivo.* This adjustment showed no changes in the viral transcript levels of

the producer cells, and no significant reduction in viral titre. In the study, expression of the transgene *in vivo* in both neurons and retinal cells was improved with this SIN vector due to the removal of transcriptional interference from the HIV-1 LTR promoter/enhancer. The study reported a deletion of the enhancer/promoter sequences in the U3 region of the 3' LTR (120-40bp ΔU3).

Similar results were obtained by Zufferey and colleagues (Zufferey *et al.*, 1999) with deletion of up to 400bp of the 3' LTR U3 region. Again, viral particle production was neither decreased nor was transduction efficiency lowered *in vitro* or *in vivo*. Furthermore, it has been demonstrated that in SIN vector transduced cells that were subsequently infected with wild type HIV-I, the vector was not mobilised (Bukovski *et al*., 1999).

The use of SIN LV significantly reduced promoter interference and also offered a potential safety advantage over traditional retroviral vectors. Specifically, upon reverse transcription, during the step of viral replication, the U3 will be duplicated in both LTRs, thus both LTRs are inactivated leaving only the very 5' end of the U3 for integrase recognition and function. Expression of the transgene will thus be controlled from an internal promoter.

If the LTR deletion is sufficient, transcription of full-length vector RNA transcripts is markedly reduced in cells transduced with a SIN vector. This further minimises the risk of generate replication competent retrovirus and there is a rather incomplete U3 enhancer sequence in the proviral 5' LTR interfering with internal, heterologous promoters. This offers additional safety advantages as it is more difficult to re-generate a wild-type parental retrovirus via recombination. With the deletion of the viral transcriptional elements from the vector, synthesis of vector RNA will depend on site of integration (Kappes & Wu, 2001).

Furthermore, the SIN design prevents transcriptional interference by the promoter/enhancer elements in the host genome, reducing the possibility of insertional activation and mutagenesis of adjacent coding sequences (such as oncogenes) either by transcription from the 3' LTR at the site of vector integration or LTR enhancer activation of a host gene promoter. In addition, to improve biosafety,

the use of SIN vectors has two additional advantages; elimination of transcriptional interference by the LTR promoter, and the possibility to create tissue-specific and inducible vectors via appropriate internal promoters, which would be difficult in the presence of non-specific transcription from the LTR promoter.

1.3.3.4 Infection by Lentiviral vectors

The nuclear import machinery recognizes the cellular matrix and docks the PIC complex to the nuclear pores, enabling the passage into the nucleus. Only a short fragment of the DNA is synthesised before the 5' end of the 5'LTR is reached and then the complex shifts to the 3'LTR of the RNA template where reverse transcriptase synthesises a new full ssDNA strand as template. The synthesis of the second ssDNA is initiated from two sites at the original (now degraded) RNA - the 3' polypurine tract (PPT) and the central PPT (cPPT), this process is decribed in detail in **Figure 1.12.**

The now dsDNA traffics towards the nucleus associated with viral proteins as linear double-stranded cDNA comprising the PIC (Herschlag & Brown 1999). Lentiviruses, unlike oncoretroviruses, rely on active transport of the PIC through the nuclear pore by the nuclear import machinery of the target cell, allowing the viral proviral genome to integrate into the host cell DNA without the necessity for cell division. The PIC traverses the nuclear membrane directly without requiring its brake down by cell division.

Integrated wild type provirus is able to be transcribed into unspliced full length copies and multiple spliced copies from the 5´LTR. Unspliced copies are packaged into new virions whilst spliced versions are used as templates for translation to produce the viral proteins necessary for this process (Pluta & Kacprzak, 2009).

A main advantage in the use of LV is therefore the easy access into the cell nucleus without the necessity for mitosis with increased probability for genome integration.

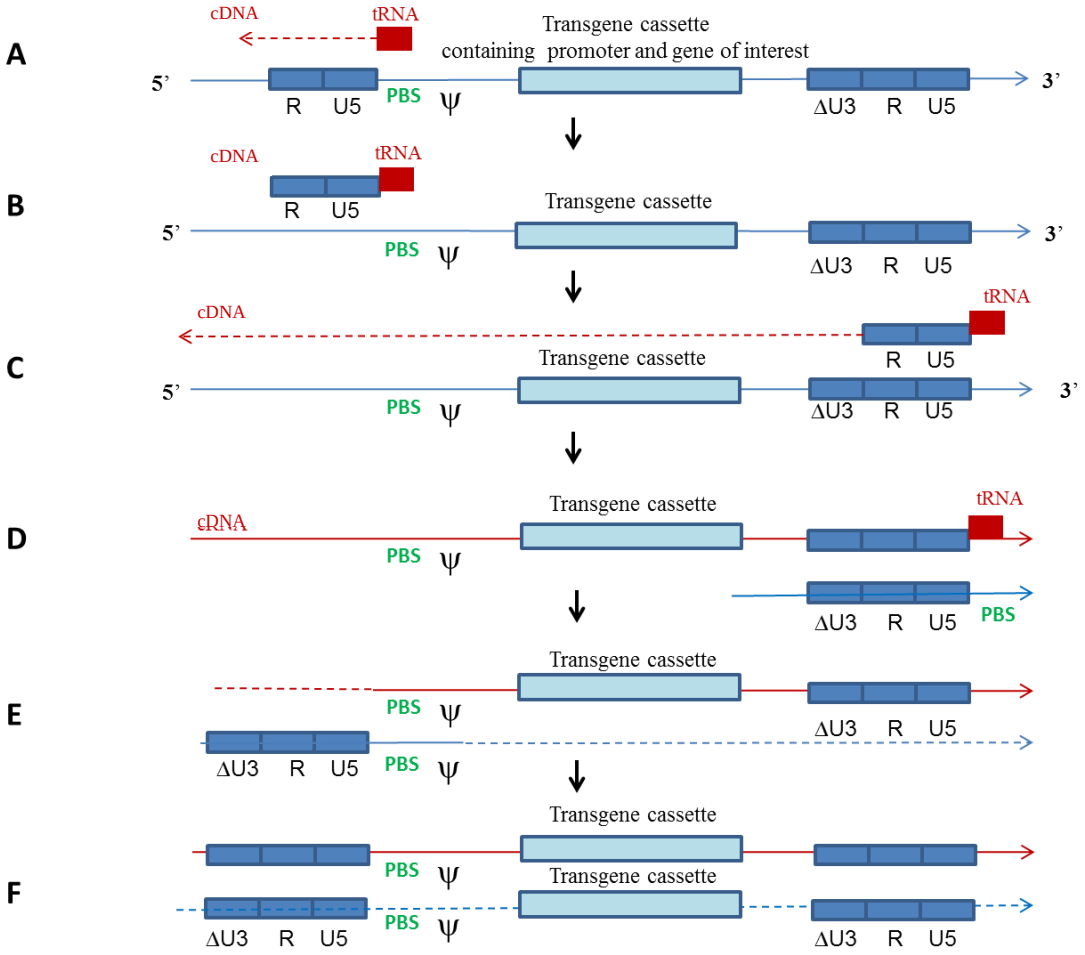


Figure 1.12 – Reverse transcription of SIN vectors. (A) cellular tRNA hybridises with the primer binding site (PBS) on the vector RNA genome and reverse transcriptase makes a DNA copy of the U5 and R regions. **(B)** The vector U5 and R regions are removed by the RNase H activity of reverse transcriptase. **(C)** The cDNA hybridises with the complementary R region at the 3' end of the vector RNA genome and the remainder of the vector genome is copied into cDNA. **(D)** Most of the remaining vector RNA genome is removed by RNase H. A second cDNA strand is extended from the remaining viral RNA and the deleted U3 region of the 3' LTR copied. **(E)** The tRNA and remaining vector RNA are removed by RNase H. The copied PBS region in the second DNA strand hybridises with the PBS region on the first strand and both strands are extended. **(F)** A double-stranded DNA copy of the retroviral vector genome is produced with 'self-inactivated' LTR regions.

1.3.3.5 Long-term tissue specific expression

For many diseases long-term and stable expression of a therapeutic transgene delivered tissue specifically is crucial. Unfortunately, progressive silencing of a stably integrated transgene over time in a cell line or *in vivo* commonly occurs. This gene silencing, commonly referred to as "position effect variegation" (PEV), is due to chromatin and DNA methylation mechanisms even though LVs preferentially integrate into open active chromatin within the target cell at the time of transduction. Therefore, LVs can be highly prone to silencing due to extensive changes in chromatin conformation.

1.3.3.6 Insertional mutagenesis

The gamma-retroviral and lentiviral vector point of insertion into the genome cannot be targeted. However, integration in 60-70% of cases is now known to occur near or within actively transcribed genes. This includes integration within or near genes involved in regulating cell growth (Imren *et al.*, 2004). As a result, problems arising from insertional mutagenesis are a possibility (Lutzko *et al.*, 2003).

Recently, the successful treatment of SCID-X1 and SCID-ADA by gene therapy has been reported (Cavazzana-Calvo *et al*., 2000; Michallet *et al*., 2000).

X-linked severe combined immunodeficiency (SCID-X1) results from mutations in the *IL2RG* gene, which encodes the common cytokine receptor gamma-chain (γc). The failure of γc signalling results in a classical phenotype characterised by the absence of T and natural killer (NK) cells, whilst B cells are present but poorly-functional (White *et al*., 2000). Unless patients are treated by haematopoietic stem cells (HSC) transplantation, they succumb to infections during the first years of life. The aim of two clinical trials using gene therapy for the treatment of this disease was immune reconstitution with autologous HSCs manipulated to express γc. The *ex vivo* procedure entailed isolating HSCs from patient and transduced with a

gammaretroviral vector expressing γc. The cells were then infused back into patients and reconstitution of the immune system was monitored over time. Overall 20 patients have been subjected to gene therapy correction (10 in Paris, 10 in London). However, 5 out of the 20 children enrolled in these two trials developed T-cell leukaemia-like conditions. In 4 out of the 5 cases this was due at least in part to the transgene activation of the LMO2 proto-oncogene, leading to aberrant, highly elevated translation of the protein, associated with childhood leukaemia (Cavazzana-Calvo *et al.* 2000). Although there is already a strong advantage for the genetically modified cells to proliferate, the activation of *LMO2* expression probably boosted this ability of the clones to the point of malignancy (Hacein-Bey-Abina *et al*., 2003a; Hacein-Bey-Abina *et al*., 2003b).

This outcome prompted the scientific community to endeavour all necessary efforts to understand how viral vectors can cause insertional mutagenesis. Several *in vivo* and *in vitro* assays have been developed to investigate and quantify this process. A protocol has been adapted from the observation that primary bone marrow cells can be immortalised by infection with gammaretroviruses (Du *et al*., 2005). This *in vitro* method is able to quantify insertional mutagenesis that leads to transformation by comparison with background levels. It has been possible to establish that the transformation frequency detected with a SIN retroviral vector is lower compared with a non-SIN counterpart and that the cellular EF1α promoter used as internal promoter in a SIN retroviral vector does not give rise to detectable transformed cells (Zychlinski *et al*., 2008; Modlich *et al*., 2006).

Until recently, very little was known concerning integration selectivity. Genome-wide studies using human primary cells are an obligatory step to accurately estimate risks. Studies performed by Wu and colleagues (Wu *et al*., 2003) confirmed that active gene vicinity appears to offer favoured sites for retroviral integration. However, it is important to refer that up to now, no other integration event affecting gene transcription leading to adverse effects has been reported in clinical trials involving retroviral vectors, which might suggest that the transgene or the target cell may have a influence in malignancy arising. There is therefore a need to achieve therapy with only 1-2 copies to minimise the risk of insertional mutagenesis.

The need for fewer vector copies per cell means that in the cases where high levels of expression per transgene copy are needed to achieve maximum therapeutic value, especially in the case of the haemoglobinopathies, it is essential to construct vectors that produce high levels of *HBB* or *HBG*. A key challenge for clinical therapies based on retroviral vectors is then to achieve stable transgene expression while minimising insertional mutagenesis whist avoiding silencing. A possible solution is the use of viral vectors with control elements that lack enhancer function (such as the UCOE; Zhang et al., 2007 & 2010) or to include elements with enhancer blocking activity into the vector to reduce the risk of oncogene activation.

Alternatively, it is also possible to consider the development of mutant LV packaging plasmids, with a mutant integrase protein, which co-transfected with transfer and envelope plasmids generate defective properties of insertion, with capacities of cell transduction but remaining in the cell nucleus as a circular dsDNA circle.

1.3.3.7 Integration-defective lentiviral vectors

As described previously retroviral and lentiviral vectors can be made integration defective (ID) by mutation of the integrase coding sequence, resulting in normal DNA synthesis but lack of integration. ID-LV genomes persist in transduced cells as either non-integrated linear or mainly as double-stranded circular genomes. These are produced during the normal replication cycle, when a proportion of reverse transcribed double stranded DNA (dsDNA) lentiviral genomes undergo either intramolecular recombination of the LTRs or intramolecular end-ligation to form closed circles containing 1 (1-LTR circle) or 2 (2-LTR circle) copies of the LTR respectively (Shank *et al.,* 1978; Engelman, 1999)

Studies demonstrating for the first time efficient and sustained gene expression *in vivo* with an integrase-deficient HIV-1 LV were conducted in $Rpe65^{rd12/rd12}$ mutant mice, a rodent model of retinitis pigmentosa (Yanez-Munoz *et al.* 2006). This study showed that high efficiency of gene transfer and expression prompted by LVs can occur *in vivo* without a requirement for vector integration. In addition, this

investigation demonstrated the potential advantage of non-integration allied with the greater transgene capacity of ID-LVs (compared to recombinant adeno-associated virus vectors) and their low immunogenicity.

This system seems to be particularly feasible for gene therapy targeting post mitotic tissues, such as muscle, liver, brain and retina. However, for target tissues constantly undergoing cell division, such as the haematopoietic system and for diseases that require early treatment for good prognosis such as Duchenne muscular dystrophy, the success of a treatment with ID-LVs is only an option if the vectors can be maintained in replicating cells. The circular forms of dsDNA ID-LV genomes persist in transduced cells as episomal plasmid-like molecules from which efficient gene expression can still take place. However, as these are unable to replicate, and thus are gradually lost from actively dividing cell populations (akin to transiently transfected plasmids, adenoviral and AAV vectors). Therefore, a means of conferring replication and retention of ID-LV is required to allow their use in mitotic stem cell populations that are the target in the haemoglobinopathies. A possible solution for this problem is the insertion of effective origins of replication (ori) and retention elements. Development of systems that employ the SV40 Ori have shown encouraging results in the development of long-term episomal maintenance in vitro, but the system requires a transforming viral gene product Large T antigen for episomal maintenance (Piechaczek *et al.,* 1999).

Despite this, the vector system has numerous applications. In two recent papers a Sleeping Beauty (SB) transposase-mediated transposition is described, which was mediated using an ID-LV (Vinc *et al.,* 2009; Staunstrup *et al.,* 2009). Unlike retroviral integration, SB transposition shows little preference for specific genomic regions and therefore a reduced risk of insertional mutagenesis. Similarly, platforms that enhance site specific integration such as zing-finger nuclease systems, used a three ID-LV system where vectors carried either one of the two necessary zinc-finger nucleases or the donor template (Lombardo *et al.,* 2007). Successful gene correction was observed at the *IL2RG* locus and site-specific integration into the CCR5 locus in 50% of transduced cultured cells and 5% of human ES cells. Initial success is,

however, still hampered by a very low efficiency in both systems (Banasik & McCray, 2009).

ID-LVs represent a class of novel viral vectors with emerging applications. However, further *in vivo* testing is required to prove the value of these techniques.

1.3.4 Success in clinical trials using viral vectors

1.3.4.1 Trials of gene therapy for X-linked severe combined immunodeficiency (SCID–X1)

X-linked severe combined immunodeficiency disease (SCID-X1) is an inherited disease with absent T lymphocyte function, due to deficiency in the common cytokine receptor γ-chain (γc; IL2RG). IL2RG is an essential component of the IL-4, 7, 9, 15 and 21 cytokine receptor complexes. A major breakthrough in gene therapy came in 2000 when successful treatment of 2 children suffering from SCID-X1 was achieved by the introduction of a normal functioning copy of the IL2RG cDNA in patient bone marrow HSC using a gammaretroviral vector based on murine Maloney leukemia virus (MLV) by an *ex vivo* delivery procedure (Hacein-Bey-Abina *et al.*, 2002; Escors & Breckpot, 2010; Cavazzana-Calvo *et al.* 2000). A similar and equally successful approach was employed in a trial in London (Gaspar *et al.*, 2004*).*

The efficacy and safety of the treatment has been demonstrated, as no evidence of replication retrovirus detected from the patients. At the end of the trials a total of 20 patients were treated; 10 in London and 10 in Paris. All patients responded well to treatment with eldest two patients now 15 years of age and in good health. However, in both clinical trials, 5 cases (4 in Paris and 1 in London) of a leukaemia-type condition directly associated to the gene therapy itself were later reported. Upon detailed analysis, the cause of these severe adverse events was found to be due to retrovirus-induced insertional mutagenesis. In four cases provirus integration within or near the *LMO-2* locus led to activation of LMO-2 causing the observed clonal acute lymphoblastic leukaemia (Hacein-Bey-Abina *et al.*, 2003; Howe *et al.*, 2008*;* Escors & Breckpot, 2010).

1.3.4.2 Gene therapy for chronic granulomatous disease (CGD)

Following a similar approach to the SCID-X1 clinical trials, the correction of X-linked chronic granulomatosis (X-CGD) was reported in 2006, again using gamma-retroviral vector encoding gp91 phox targeting bone marrow HSC (Ott *et al.,* 2006). Clinical efficacy was observed, together with clonal amplification of corrected cells, probably as a result of insertional activation of the MDS1-EVI1, PRDM16 and SETBP1 genes (Stein *et al.,* 2010). In this trial a moderate dosage of busulfan (8 mg/kg) was given prior to reinfusion of the transduced mobilised HSC (Ott *et al.,* 2006). The two adult males enrolled in the trial showed relatively high levels of corrected leukocytes in peripheral blood (~20%) within the first months after the gene therapy procedure, which rose to as high as 80% over the first year. Vector integration sites revealed a highly restricted pattern, with the majority of vector integrants in the engrafted stem cells being near one of a few genes known to be involved in myeloid cell proliferation (*MDS-1, PRDM16* or *SETBP1*). These two patients went on to develop a benign form of myelodysplasia, a pre-leukemic condition. After initial success with both patients showing clearance of chronic pre-existing infections, one patient unexpectedly developed an acute infection from which they subsequently died. Examination revealed that although neutrophil counts had been maintained at normal levels therapeutic transgene function had been lost in both patients. The underlying mechanisms for the oligoclonal expansion and progression to myelodysplasia are not fully elucidated, but the retroviral vector used contained the gp91phox therapeutic gene under control of the SFFV LTR, which possesses potent enhancer activity in myeloid progenitor cells and it is thought that this led to *trans*-activation of genes that promote myeloid cell proliferation. Molecular analysis showed that the promoter region of the SFFV element had become progressively methylated and silenced leading to observed failure of the therapy. However, the enhancer region of SFFV had remained unmethylated and active, thus retaining its ability to induce expression of the *MDS-1, PRDM16* or *SETBP1* promoters near the site of retroviral vector integration. The use of HSC target cells that that had been mobilised from bone marrow into the circulation by the myeloid growth factor G-CSF may have presented a large number of transcriptionally active myeloid-proliferative

genes as integration targets. As with the SCID-X1 trials, it will be vital to understand how to avoid this unwanted complication while retaining the clear-cut clinical benefits that can be attained. In addition, the problem of therapeutic gene silencing that ultimately led to treatment failure (Ott *et al*., 2006; Stein *et al*., 2010) needs to be solved. In this regard very encouraging results have recently been reported with an A2UCOE-MRP8-gp91phox transgene that combines both stability and neutrophil specific expression (Brendel C *et al.,* 2011)

1.3.4.3 Gene therapy for Leber's Congenital Amaurosis (LCA)

Encouraging improvements of vision were reported in a genetic eye disorder leading to early childhood blindness - Leber's Congenital Amaurosis (LCA) - using AAV-2 vector mediated gene therapy (Bainbridge *et al.,* 2006, Alexander *et al.* 2007). Approximately 10% of LCA cases are defective in the *RPE65* gene, which is required for normal retinal cycling of vitamin A. Thus, gene addition gene therapy was carried out in this study by using AAV2 as a vector to replace *RPE65* in retinal pigment epithelium (RPE). In pre-clinical studies in dogs, this treatment improved visual function, which was maintained for more than three years (Bainbridge *et al.,* 2006). The results of a phase I clinical trial involving three young adult patients with early-onset, severe retinal dystrophy by mutation in RPE65 have been reported (Bainbridge *et al.,* 2006). There was no immune response reported but no clinically significant improvement in visual acuity and retinal response to flash light or electroretinography in all three patients. However, one of the patients has improved in retinal function and visual mobility in low light condition (Bainbridge *et al.,* 2006). The safe outcome from this initial low dose cohort of patients has led to approval for vector dose escalation to levels expected to give consistent therapeutic effects.

A similar clinical trial for LCA has also been conducted in the USA (Maguire *et al.,* 2008). An AAV2 vector with an *RPE65* therapeutic gene was used in this trial. Visual perception was observed to markedly improve in a dimly lit environment for all three patients treated. The improvement in pupillary light reflexes confirmed the increase in retinal sensitivity and visual acuity (Maguire *et al.,* 2008). Importantly, no

adverse effects were reported, including response to RPE65 protein. The safety and efficacy of the vector has persisted through to at least 1.5 years post-injection (Simonelli *et al.*, 2010).

Participants reported baseline vision of about 20/200 with statistically significant visual acuity improvement and significant increases in light perception. Light sensitivity measurements were also reported to increase, with the vector-treated eyes showing a greater pupillary response to light than at baseline and to the previously better seeing contralateral eye - Injections were performed on the patient's worse seeing eye. The positive initial findings with respect to both efficacy and safety using AAV-mediated gene therapy for LCA are exciting and encouraging (Chung & Traboulsi, 2009; Simonelli *et al.*, 2010).

1.3.4.4 First human gene therapy trial using lentiviral vectors – gene therapy for X-linked adrenoleukodystrophy (ALD)

More recently, the first clinical gene therapy trial employing a lentiviral vector for an inherited disease, X-linked adrenoleukodystrophy (X-ALD), has been reported (Cartier *et al.*, 2009). X-ALD is a severe brain demyelinating disease caused by mutation of *ABCD1* encoding the ALD protein, which is an adenosine triphosphate-binding cassette transporter in the membrane of peroxisomes (Cartier *et al.*, 2009). The lentiviral vector expressed a wild-type *ABCD1* cDNA under the control of the myeloproliferatuve sarcoma virus enhancer (MND), and was used to genetically correct autologous cytokine-mobilised peripheral blood $CD34^+$ cells *ex vivo* and transplanted into three young ALD patients after pre-transplant conditioning with myeloablative doses of busulfan and cyclophosphamide. This gene therapy approach has been found to stop progression of the disease with the same efficacy as transplant of bone marrow HSC. The therapeutic effect has been achieved to correct the phenotype of ALD for 24 to 30 months (to date), with 9 to 14% of granulocytes, monocytes and T and B-cells expressing the ALD protein.

Notably, this study demonstrated that lentiviral vectors can mediate significant levels of gene transfer into human HSCs.

1.4 Gene Therapy for the Haemoglobinopathies – past, present and future

1.4.1 Review of gene therapy for haemoglobinopathies

The transfer of a regulated *HBB* transgene into autologous bone marrow HSC is a highly attractive treatment. The strategy, simple in principle, raises major challenges in terms of controlling expression of the globin transgene, which ideally should be erythroid specific, differentiation and stage restricted, position independent and sustained over time at a sufficiently high level to be of therapeutic benefit. The gene therapy approach would resolve the problem of finding an HSC donor and eliminate the risk of graft-versus-host disease and graft rejection associated with allogeneic HSC (Quek & Thein, 2007).

Achieving regulated *HBB* expression has represented a tremendous obstacle in the past 20 years. Early attempts using gamma-retroviruses showed tissue specificity but low and variable levels of *HBB* expression in murine bone marrow chimeras due to vector instability (Cone *et al.,* 1987; Dzierzak *et al.,* 1988). A new hope came with the discovery of the LCR. Initial efforts to incorporate βLCR sub-fragments into oncoretroviral vectors significantly increased expression levels in murine erythroleukemia (MEL) cells but failed to abolish positional variability of expression (Sadelain *et al.,* 1995). This finding suggested that a minimal LCR comprising the juxtaposed core elements did not provide full LCR function, but rather acted like an erythroid specific enhancer.

The later incorporation of larger βLCR segments into retroviral vectors proved to be problematic, leading to vector instability and considerable genomic rearrangements and very low viral titre.

Tests with AAV vectors (Einerhand *et al.,* 1995) were soon discarded as these have a cloning capacity of only 4.7kb and are therefore unable to accommodate the required transgene size – with only the ‘core’ elements of the βLCR capable of being inserted and therefore these failed to overcome positional variability of expression in

murine models. Furthermore, AAV vectors were subsequently found not to integrate efficiently into the target cell genome, therefore providing only short-term expression in dividing cells.

1.4.2 Targeting HSCs with lentiviral vectors - milestones

Soon became evident that lentiviral vectors would be of great use in the development of erythroid-specific gene therapy vectors. LVs show distinct advantages: they transduce non-dividing cells, achieve relative genomic stability, are generally considered safe vectors with a self inactivating (SIN) design that results in the removal of the viral LTR upon proviral integration, and have a relatively large packaging capacity (up to 8kb). Indeed, LVs remain to date the optimal vector where stable integration into the target cell genome is required, and thus hold the greatest promises of therapeutic success. As they have proved to stably incorporate larger DNA inserts without rearrangements, this allows the incorporation of larger βLCR fragments and chromatin insulator elements to be used in an effort to reduce position effects, with the ability to obtain sustained therapeutic levels of expression.

The breakthrough which changed the general negative thoughts surrounding gene therapy for the haemoglobinopathies occurred when the laboratory of Michele Sadelain showed definitive proof of principle that an LV harbouring an optimised combination of proximal and distal *HBB* transcriptional control elements, could yield therapeutic levels of HBB expression and completely rescued the disease condition in in the *th3* severe thalassaemia *intermedia* mouse model system (May *et al.*, 2000). They reported a lentiviral cassette encoding an *HBB* containing its own promoter including both introns, the intragenic enhancer (located in intron 2) and 3' proximal enhancer (constituting an HBB-mini-gene) together with larger fragments of the βLCR sites HS2, HS3 and HS4. This vector was termed TNS9 (**Figure 1.13**).

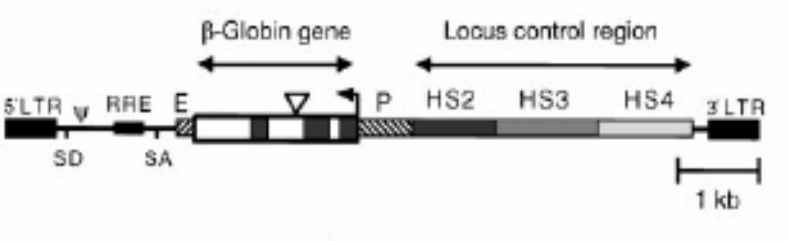


Figure 1.13 - Diagram of the TNS9 globin lentiviral vector. The image shows the cassette incorporated into the transfer plasmid. The βLCR sites (HS2, HS3, HS4) are located 1kb from the 3´LTR. Exons and introns of *HBB* are represented by filled and open boxes respectively. Note non-SIN vector configuration. (Image taken from May *et al.*, 2000).

The TNS9 vector has been shown to drive erythroid-specific expression of *HBB* at therapeutic levels in thalassaemic mice (May *et al.,* 2000). At four months post-transplantation, mice harboured on average 0.5-1 vector copy per peripheral blood cell and showed haemoglobin levels of 11-13g/dl in comparison with 8-8.5g/dl in age-matched controls. The carefully considered combination of βLCR HS sites together with *HBB* itself in this vector enhanced expression beyond levels previously achieved using arrayed minimal core elements (Rivella & Sadelain, 1998), yielding expression within a therapeutic range and at low copy number. This led to phenotypic improvement in heterozygous *th3* β-thalassaemic mice: delivery of *ex vivo* transduced autologous bone marrow cells maintained transgene expression for over a 40-week period (May *et al.*, 2002). However, it should be noted that the TNS9 is of a non-SIN configuration and thus integrated proviral genomes of this vector will harbour intact 5' and 3' LTR elements. Therefore, despite its apparent efficacy on safety grounds TNS9 cannot be employed within a clinical setting.

The therapeutic levels of *HBB* expression achieved under this system soon paved the way for the development of other HIV-1 based LVs for gene therapy of the haemoglobinopathies. However, the TNS9 vector required at least one copy of the vector in every stem cell to correct severe murine β-thalassaemia requiring most or all of the HSCs to be modified suggested that inconsistent globin expression could be occurring. A different study performed simultaneously (Rivella *et al.,* 2003) further supported this concern: the administration of TNS9 to a mouse model with severe

anaemia in which mice engrafted with β-globin–null (Hbb*th3/th3*) foetal liver cells succumb to ineffective erythropoiesis within 60 days, only 1 of 6 animals demonstrated a relatively high level of haemoglobin following the gene transfer procedure despite nearly all the HSCs being genetically modified. The remaining animals remained severely anaemic with a thalassaemia *intermedia* phenotype.

1.4.3 Correction of sickle cell anaemia mouse models

Leboulch and colleagues were the first to demonstrate haematological correction and diminished end organ damage in two murine sickle cell model systems using lentiviral-mediated HSC gene transfer of an anti-sickling variant of the human HBB chain (Pawliuk *et al.,* 2001). The authors reported a structural optimisation of the HBB/LCR lentiviral cassette, which resulted in high viral titres capable of engendering multiple events of chromosomal integration per HSC (average vector copy number of 3). This led to a balanced expression at a sufficiently high and homogeneous level to provide expression similar to that observed in asymptomatic carriers (Pawliuk *et al.,* 2001). This result was later confirmed by others (Levasseur *et al.,* 2003) with a different sickle cell mouse model and a slightly different anti-sickling *HBB* vector cassette. Phenotypic improvement with an average vector copy of 2.2 resulted in haemoglobin tetramers incorporating the transgene *HBB* chain at a level of 20% of the total haemoglobin for at least a period of 7 months.

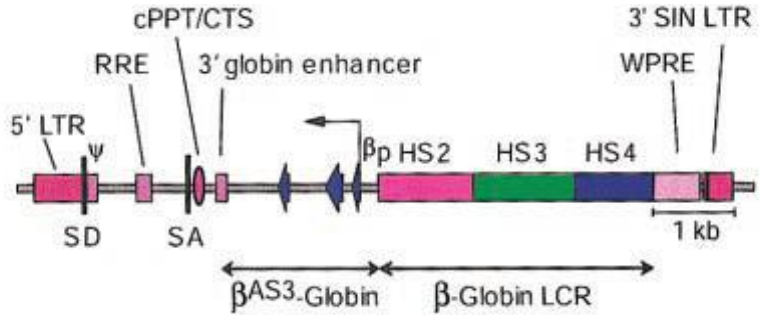


Figure 1.14 - The lentiviral/anti-sickling HBB expression construct designed by the Townes lab (Levasseur *et al.,* 2003). HS2(1203 bp), HS3 (1213 bp), and HS4 (954 bp) sequences of the βLCR, the 3' globin enhancer, the 266bp *HBB* promoter and the β anti-sicking β^{AS3}-globin gene are shown.

1.4.4 *HBG-A-based* lentiviral vectors

Parallel efforts have also been made with the construction of LVs designed to improve foetal haemoglobin levels (*HBGy-Ay*). *HBGy-A* naturally switch off at birth but its known to ameliorate the clinical severity of both β-thalassaemia and sickle cell anaemia in adults in cases where hereditary persistence of foetal haemoglobin (HPFH) has been co-inherited with the severe haemoglobinopathy. In 2004 the Persons laboratory developed a *HBG* LV, having more extensive LCR-derived regulatory sequences (**Figure 1.15**). This vector was shown to be less susceptible to position effects, was able to produce higher and more consistent levels of HbF protein with one vector copy per cell being curative (Hanawa *et al.,* 2004). More recently, this same laboratory has made an additional significant contribution for HSCs *in vivo* selection. A foetal Haemoglobin/ methylguanine-methyltrasferase (MGMT) dual gene LV could be used effectively for *in vivo* selection of HSCs following cytotoxic drug administration, resulting in large increases in HBB expressing cells and resolution of anaemia in a mouse model of β-thalassaemia (Zhao *et al.,* 2009). This value compares favourably with their single-gene vector (Hanawa *et al.,* 2004 - 2 g/dL per vector

copy) and also to those observed by others using HBB LVs and similar mouse model systems.

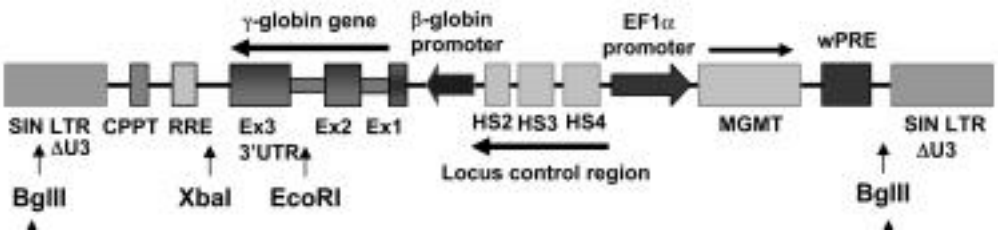


Figure 1.15 - Schematic diagram of the γ-globin/MGMT lentiviral vector used by Persons lab. (*Image taken from* Zhao *et al.*, 2009).

1.4.5 Modifications to the initial TNS9 cassette

In 2004 the Lebouch laboratory developed an LV containing an anti-sickling *HBB*$^{A\text{-}T97G}$ expression cassette that included a 2.7kb region of the βLCR HS2, HS3 and HS4 sites plus mini-*HBB* (**Figure 1.16**). This was used to transduce normal human haematopoietic stem cells cells and was capable of establishing haematopoiesis in immunodeficient mice. Approximately 80% of the HSCs were genetically modified, resulting in chimeric haemoglobin molecules incorporating human HBB comprising about 21% of total haemoglobin (Imren *et al.*, 2004).

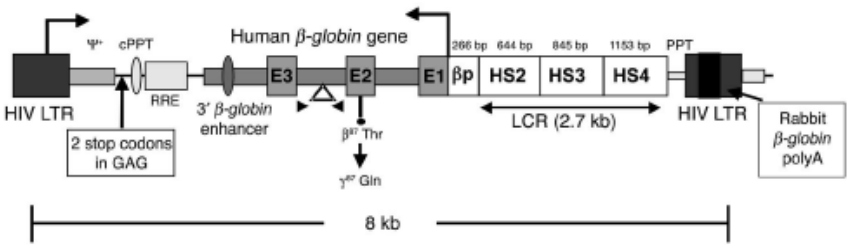


Figure 1.16 - Schematic of the $\beta^{A\text{-}T97G}$ LV (Imren *et al.*, 2004). It includes HS2, HS3 and HS4 βLCR sites regulating expression of the *HBB* transgene placed in reverse orientation.

The results of this study showed that efficient transduction of HSCs from normal cord blood engrafted into an immunodeficient mouse model expressed *HBB* in their erythroid progeny at sufficient levels to improve SCA and β-thalassaemia phenotypes. However, even with this vector, multiple vector copies are needed per cell for a sufficiently high level of transgene expression and risks of insertional mutagenesis have highlighted the need to achieve therapeutic goals with minimal copies of proviral integrants per cell. Also, it is evident in this study that proviral integration occurs frequently within introns of actively transcribed genes that play crucial regulatory rules in haematopoiesis, which once again raises questions on the long-term safety of this vector (von Kalle *et al.*, 2004).

The laboratory of Punam Malik (Puthenveetil *et al.*, 2004) showed the first evidence of effective genetic correction of β-thalassaemia phenotype using an *in vitro* and an *in vivo* xenograft model of human erythropoiesis with a cHS4-insulated globin vector, BG-I (**Figure 1.17**). The BG-I vector was used to transduce $CD34^+$ cells from four patients with thalassaemia-*major*. Genetically modified thalassaemia *major* $CD34^+$ cells were then transplanted into splenectomised β2Mnull-NOD-SCID (immunodeficient) mice and 3-4 months later, effective production of human glycophorin A–expressing erythroid cells were observed in the circulation. There was reversal of ineffective erythropoiesis and effective RBC production.

Figure 1.17 - The BGI vector (Puthenveetil *et al.*, 2004). SIN lentiviral backbone containing a 1.2-kb cHS4 insulator (cHS4-I) element inserted to replace the 398-bp

U3 promoter and/or enhancer deletion, a 3.1kb human βLCR region consisting of hypersensitive sites (HS) 2, 3, and 4; HBB promoter (254 bp); the *HBB* gene (with a 372-bp deletion in IVS2); and the 3' enhancer (3'E) were cloned in reverse orientation to the viral transcriptional unit.

The use of the chicken HS4 element flanking an *HBB* cassette contributes to reduced variability of transgene expression, thus partially protecting the provirus from chromosomal position effects, and resulting in an overall 2 to 4 fold increase in average expression (Arumugam *et al.,* 2007). This design, although promising, results in reduced viral titre making it difficult to be produced large-scale as required for clinical trials.

In 2007, the Sadelain laboratory published a report claiming that the addition of HS1 of the βLCR in their basic TNS9 LV design did not improve *HBB* expression, but the addition of HS4 in addition to HS2 and HS3 improved the expression from 27 to 41% per vector copy. However, the results of this report should be analysed with care with regards to the vector design; one familiar with the βLCR looping mechanism and chromatin hub formation cannot but wander if this result is, in fact, based on the interaction of HS1 or a merely a consequence of the increased physiological distance of the βLCR form the *HBB* in the vector improving ACH formation and enhanced gene expression.

An *HBB* transgene under control of a minimal 2.7-Kb HS2 and HS3 βLCR element combination has been shown to produce efficacious amounts of HBB protein (Miccio *et al.,* 2008). This report showed sufficient data to demonstrate that only these 2 βLCR elements are sufficient for stable transgenic expression in tissue culture erythroid-specific cells. Studies performed with this "GLOBE" LV have shown it capable of correcting of β-thalassaemia in mouse model systems achieved by *in vivo* selection of genetically corrected erythroblast progenitors, differentiating from a relatively limited number of transduced HSCs, providing high transcription levels of

HBB that rescue the β-thalassaemia *intermedia* and major phenotype (Miccio *et al.*, 2008).

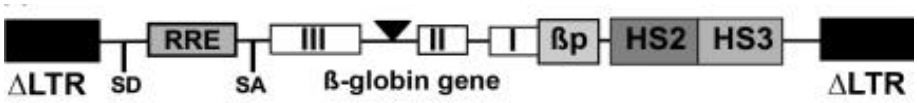


Figure 1.18 - The GLOBE vector (Miccio *et al.*, 2008). Human β-globin gene, β-globin promoter (βp), and DNase I-hypersensitive sites HS2 and HS3 from βLCR are shown.

The GLOBE vector (**Figure 1.18**) has also recently been shown to correct the globin chain imbalance following transduction of $CD34^+$ cells from β-thalassaemia patients (Roselli *et al.*, 2010). Firstly, the authors demonstrate at the *ex vivo* manipulation required for successful transduction do not affect the viability of CD34+ cells and proceeded to demonstrate that at low copy number lentiviral transduction of CD34+ is capable of correcting β-thalassaemia and restore erythropoiesis upon differentiation *in vitro.* Finally, the authors analysed vector integration sites and their proximity to proto-oncogenes, reporting that, together with a CMV-GFP vectors, the therapeutic vector integrates preferably within transcriptionally active genes, with no bias towards integration in proximity of oncogenes. More importantly, despite integration within or near active genes GLOBE did not give rise to increased host gene expression suggesting it has a low insertional mutagenesis potential (Roselli *et al.*, 2010).

More recently it has also been published the first clinical trial for β-thalassaemia involving LVs employing BM HSC *ex vivo* procedure, coupled with full myelo-ablative conditioning (Cavazzana-Calvo *el al*., 2010). An 18year old patient with transfusion dependent HbE/β^0-thalassaemia was treated with an LV containing a

functional *HBB*. The patient underwent a high dose of chemotherapy to eliminate all diseased HSCs before being administered with their own genetically modified HSCs. The HSCs containing the transferred *HBB* gradually gave rise to healthy blood cells and the levels of normal HBB protein increased with 10-20% of reconstituted HSCs containing the transferred *HBB*, although only 1/3 of Hb was from the LV with the remaining 2/3 coming from high endogenous HbF (1/3) and HbE (1/3). One year after the treatment the patient no longer needed blood transfusions.

Furthermore, 50% of genetically modified cells and consequent increased levels of *HBB* seen in this patient, were a result of clonal dominant expansion of a single myeloid progenitor with a site of LV integration within the *HMGA2* gene resulting in expression of an exons 1-3 truncated product at markedly higher levels. *HMGA2* encodes a protein regulating gene expression and its disturbance has been reported as potential oncogenic (Fusco *et al.*, 2007). Therefore, regardless of the author's claims of therapeutic success, this study reports a clear example of a fortuitous insertional mutagenesis event contributing to therapeutic. An outcome whose longer-term consequences are unknown.

1.4.6 Gene therapy for the haemoglobinopathies – future challenges

Results from several patients in recent gene therapy trials published to date in which LVs were used to transduce HSCs indicate that although clinical success can be achieved, the overall levels of engraftment with genetically modified HSCs that are currently possible are likely to be in the 10-15% range. While these levels may be sufficient to procure clinical improvement and transfusion independence in patients with thalassaemias, higher levels of gene transfer will be needed to cure patients with thalassaemia major. Nevertheless, collectively, these data supports the conclusion that *HBB* transgene expression can be reasonably regulated in the progeny of LV transduced stem cells, although position effects, transgene silence leading to variegation are limitations still not fully overcome.

1.5 In *utero* Gene Therapy

In utero gene therapy is the administration of a given vector cassette pre-natally in order to provide early therapeutic intervention and/or ameliorate a genetic defect, and thereby provide permanent somatic gene correction before birth.

1.5.1 "Proof of principle", landmarks and progress of *in utero* gene therapy

The first proof-of-principle came in 1990 (*Maria Hatzoglou et al.*, 1990). In this study the authors report that a replication-incompetent retroviral vector injected intraperitoneally into foetal rats *in utero* mediated expression of human growth hormone. This experiment paved the way and provided substantial grounds for *in utero* gene therapy with the following years seeing a large number of *in utero* marker gene experiments using several different vectors, different animal models, with successful gene delivery to various foetal tissues.

The first report of a curative gene therapy protocol *in utero* involved injection of an adenoviral vector into the amniotic fluid of CFTR knockout (KO) mice (Larson *et al.*, 1997). The injection resulted in dramatic improvement in survival of KO mice, even though CFTR expression was very limited. This report was heavily criticised in the literature at the time as relatively few data were presented with intermediate end points, such as specific mechanisms of longer-term effects of transient CFTR messenger RNA (mRNA) or protein expression or electrophysiologic correction of cAMP-activated chloride transport (Flotte, 2008). Also, this study has not been reproduced independently since.

In 2003 was performed a direct injection of an LV expressing the human bilirubin UDP-glucuronyltransferase (UGT1A1) gene under control of the

phosphoglycerate kinase promoter into the liver of the Gunn rat foetus (*Seppen et al.*, 2003*)*. The Gunn rat is a model for the very rare autosomal recessive human condition Crigler Najjar disease Type I (CN I), with bilirubin UDP-glucuronosyltransferase deficiency caused by mutations in the gene encoding UGT1A1 on human chromosome 2. The disease manifestations include neurotoxicity caused by toxic levels of bilirubin accumulating in the patient's blood leading to severe brain damage shortly after birth, and results in infant death in the absence of continual phototherapy and liver transplantation.

Direct pre-natal intra-hepatic injection of an LV delivering the corrective transgene resulted in a significant reduction (45%) of bilirubin levels for the 1 year duration of the study (Seppen *et al.*, 2003). In patients with CN I, such reduction would convert severe CN I disease to the milder CN II form. However, the absolute levels of unconjugated bilirubin observed in rats and humans are not comparable, and it is therefore not clear if the protein's expression levels would be sufficient for a therapeutic effect in man.

Successful restoration of rhodoposin synthesis and electrophysiologically measured visual functions in a mouse model of Leber's Congenital Amaurosis (LCA), a congenital retinal blindness caused by mutations in the *RPE65* gene has also been reported (Dejneka *et al.*, 2004). Subretinal application of an AAV2 vector expressing human RPE65 *in utero* resulted in efficient transduction of retinal pigment epithelium, restoration of visual function, and measurable rhodopsin. This was achieved in two of the 13 pups that survived to adulthood. Limited numbers are most likely due to technical difficulties of topical gene delivery to the retina. It is also noteworthy that RPE65 expression could be observed after 5/6 months. The retinal epithelium is virtually non-dividing and the AVV2 vector was able therefore to persist episomally.

Further successful applications *in utero* using an AAV vector have been reported (Rucker *et al*., 2004). Intraperitoneal *in utero* application of an AAV2 vector expressing the lysosomal enzyme acid alpha glucosidase (GAA) was delivered to foetuses of a mouse model of Pompe's disease. The transgene restored enzyme levels in the diaphragm of these animals, prevented glycogen accumulation in this muscle and restored diaphragm contractility up to 6 months postpartum.

More recently Waddington and colleagues has led the field of *in utero* gene therapy applications. In 2004 they performed an *in utero* administration of an LV via the foetal vitelline yolk sac vessel, a route that delivers a large proportion of the vector directly to the liver. Therapeutic levels of human Factor IX of between 18 and 32% of normal values were obtained which resulted in permanent amelioration of the bleeding disorder in Factor IX knocked out (KO) mice (a model for human haemophilia B). The KO mice have no functional Factor IX, and restoration of levels to 5% of normal corresponds in humans to a mild haemophiliac phenotype and 40% to phenotypic cure. Determination of Factor IX blood levels allowed the monitoring of transgene expression in the individual treated animals over their lifetime. Importantly, animals were kept alive for a long observation period and no antibodies against the human protein were found even after adjuvant stimulation, whereas strong antibody responses were observed after the same challenge on haemophiliac mice not treated *in utero*. In liver biopsies, human factor IX expression was detected in groups of neighbouring cells suggesting clonal propagation of human Factor IX-expressing foetal cells with progenitor function (Waddington *et al.*, 2004).

1.5.2 *In utero* gene therapy over adult gene therapy

The approach is based on the hypothesis that gene therapy performed *in utero* may prevent irreversible modifications that are likely to occur during foetal development, and avoid early onset organ damage. A pre-natal intervention avoids the development and progress of severe manifestations. Additionally, the foetal system is especially susceptible for successful gene therapy due to its immunologic naiveté. Adult gene therapy often causes an immune response that restricts transgene expression. *In utero* gene therapy takes advantage of the low predisposition for immune response and the transgenic protein would be recognised as "self" in the early stages of gene transfer in the foetal period and avoid an immune response

An *in utero* approach also takes advantage of the high proliferative status and expansion of foetal stem cells, which may result in gene expression in a large proportion of daughter cells. The relatively small amount of genetically engineered stem cells required, together with the administration of a relatively small amount of vector could result in much higher adult expression levels than what would be achieved by adult vector administration - highly attractive conditions within a gene therapy context, where the mass production of vector/stem cell numbers are limiting factors (Waddington *et al., 2004;* Coutelle *et al.*, 2003).

Although foetal gene therapy will not replace post-natal gene therapy, it constitutes a preventive approach to otherwise incurable diseases and would therefore be most effectively conducted in conjunction with pre-natal screening programmes.

1.5.*3 In Utero* procedures

1.5.3.1 The search for the optimal vector

The ideal vector for foetal gene therapy would introduce an appropriately transcriptionally regulated gene into the cells of all organs relevant to the genetic disorder by a single safe application (David & Peebles, 2008).

To date, none of the current vector systems are capable of successfully meeting these criteria, although many have characteristics that can be beneficial for a foetal approach. Lentiviruses, specifically, appear as ideal candidates in part due to the possibility of pseudotyping with different envelope proteins, allowing gene transfer to be more restricted to specific tissues and the possibility of high titre production.

Vector selection to minimise the risk of oncogenesis is crucial for *in utero* gene therapy. It has been reported that equine infectious anaemia virus (EIAV)-based lentiviral vectors have higher frequency to cause liver tumours in mouse models, while no tumour was observed from SIN HIV-based vectors (Themis *et al*., 2005). These studies are important since they suggest that liver tumour development is not caused by either injection procedure or the expression of the transgene alone, but just the vector itself (Themis *et al*., 2005).

1.5.3.2 Routes of administration

Development in vector technology has been accompanied with the development of minimally invasive methods of vector delivery. Studies in sheep have adapted ultrasound guided injection techniques. Ultrasound guidance has also been used in non-human primates to deliver gene therapy into the amniotic cavity for direct injection of the lung and liver parenchyma (David & Peebles, 2008). In mice, several

routes of application have been considered, mainly depending on the target disease; Direct injection into muscle and ocular tissues have been performed, intraperitoneal injections, via the portal vein; but is the injection of therapeutic LV into umbilical or yolk sac blood vessel, allowing the virus to gain efficient access to the foetal liver via the circulation that holds promise for people with haemoglobinopathies.

The foetus is surrounded by an interior amniotic membrane and nutrients are supplied through umbilical blood vessels in human, or the yolk sac vessels from the parietal yolk sac in mice. The foetal liver is the major haematopoietic organ and therefore the site of HSC residence during gestation. In principle it is an excellent target for gene therapy as it could provide maximum output of systemically required transgenic proteins such as Factor VIII and Factor IX, in haemophilia. LVs delivered *in utero* via the yolk sac blood vessels are able to transduce not only hepatocytes but also HSC in foetal liver before these cells migrate to bone marrow.

Many of the techniques used in animal models have successfully been used in human medicine. Fibre-optics and ultrasound guidance is routinely used for foetal surgery/direct access to the human foetal circulation. Intrauterine blood transfusions are regularly performed by injection into the intrahepatic umbilical vein and intraperitoneal injection of foetal liver stem cells has been achieved in humans from 14 weeks of gestation (Waddington *et al.*, 2005).

1.5.3.3 Time of intervention

Studies *in vivo* have proved that effective expression of a given transgene can be greatly compromised by the gestational time of delivery. The "optimal delivery period" is then variable, depending on the animal model used and route of administration of the vector.

Studies performed with mice tend to take advantage of the immune response, aiming vector administration at the end of the second gestational trimester.
Clinical data (Coutelle, 2008) has shown that vector administration before the second trimester of gestation - a "pre-immune" stage of development - can avoid the immediate inflammatory and immune response to a transgenic protein and induce immune tolerance to the transgene vector. It has been suggested that the vectors can be introduced into the human foetus before the second trimester to avoid inflammatory and immune responses and increase tolerance to transgenic protein.

The relevant time windows for the different application routes in humans remains to be established with respect to technical feasibility, foetal physiology and the development of the foetal immune system (David & Peebles, 2008). For instance, the human immune system develops from 12 to 14 weeks of gestation, with T lymphocyte maturation observed at this stage. Thus, it might be necessary to deliver a transgene before this gestational time, which greatly limits the routes of application. Experiments in non-human primates are likely to prove very useful test systems prior to clinical trials.

1.5.4 Problems associated with an *in utero* approach: ethical and safety concerns

Long-term adverse effects from pre-natal gene therapy are still not proven and unknown and it is extremely important to access the potential risks of *in utero* gene therapy thoroughly, both from a scientific, and ethical point of view. Most fears revolve around the acute toxicity and immunogenicity of the vector transgene, the possibility of germ line transmission and the danger of insertional mutagenesis/oncogenesis - some of the general risks associated with gene therapy.

One major concern regarding an *in utero* approach is the increased risk of inadvertent germ line transmission of the integrated transgene sequence. Vector integration into germ cells, if it occurs, is likely to be random, and such an event could potentially affect progeny conceived from such a germ cell.

Studies suggest that if vector administration is performed around the 2nd trimester of pregnancy, germ cells are by then well compartmentalized in their definitive organs, and could therefore be reached only via the bloodstream. It's true that vector sequences have been detected in the gonads of *in utero* injected animals, but transgene expression has never been found in purified sperms or in the offspring of these animals (Waddington *et al.*, 2005).

However, evidence for lentiviral transduction of a subpopulation of gonadal cells isolated by laser capture micro=dissection was recently reported to occur in female rhesus monkey foetuses after intraperitoneal vector administration (Lee *et al.* ,2005).

When a vector enters the foetal circulation there is a risk of it crossing the placental barrier into the maternal bloodstream. There is wide variability in placental structure among different species and only the nature of the maternal barrier will determine the likelihood of this event. Nevertheless, rodents and humans share the same class and the most permissive type of placenta (hemochorial) whereas sheep have a much less permissive (synepitheliochorial, small placentomes) placenta type, which suggests that studies on these animals, although informative, might not be physiologically relevant to humans.

The calculated frequency of naturally occurring endogenous insertional mutations in humans of about 1 in 8 individuals is substantially higher than the suggested upper tolerable limit of 1 insertion event per 6000 sperm due to a gene delivery protocol (Kazazian, 1999).

Gene therapy is not the only intragenic protocol likely to induce germline modifications. It is well known and documented such events happen with exposure to UV light and high doses of chemotherapy, to name but a few. As often with new therapeutics, only when they become efficient are potential side effects and hazards observed. Therefore, ethical concerns in relation to the balance of potential benefit and harm of an *in utero* approach are entitled to exist. The ethics of foetal somatic gene therapy have been reviewed (Wagner *et al.*, 2009). Whether or not a given disease is given approval for intrauterine intervention is a subject of criterion studies. It is important to reiterate that, to date, there is no evidence for germ line transmission by *in utero* gene therapy and therefore, the scientific community shall accept the risk/benefit associated with the procedure and advance with care performing appropriate animal experimentation with long-term post-natal follow-up before any clinical trial is approved.

1.5.5 Targeting HSCs: *in utero* gene therapy for the haemoglobinopathies

Success of adult gene therapy for genetically inherited diseases using stem cells has been hampered mainly by inefficient transduction and the inability to achieve long-term expression of therapeutic genes,

Current clinical protocols of gene therapy in patients with genetic diseases are based on *ex vivo* transduction of lymphocyte or hematopoietic stem/progenitor cells from cord blood or bone marrow, followed by autologous transplantation of these engineered cells back into the patient. Initial trials showed the feasibility and safety of gene therapy using cord blood cells in patients with ADA-deficiency,

For gene therapy to hematopoietic cells in the foetus, two different strategies are being pursued in parallel:

1.5.5.1 *Ex vivo* transduction of HSC

The approach consists in targeting autologous self-renewable HSCs from the bone marrow for defect gene correction or replacement by *ex vivo*, and then re-transplanting the genetic modified HSC back to patients/models. This methodology is especially attractive for the treatment of haematopoietic disorders caused by single gene defect, such as, β-thalassaemia and sickle cell disease (Persons, 2009).

A successful *ex vivo* strategy requires harvest of an adequate number of foetal HSC in the first trimester, efficient gene transfer to harvested cells *in vitro*, followed by re-infusion with or without expanding the transduced stem cells. A limiting factor of this technique is the narrow window in which the whole procedure must be performed (confirming genetic diagnosis, perform HSC harvest and *in vitro* transduction, and re-instilling transduced cells into the foetus).

The major advantage of this method is that the risk of germ line transduction is avoided when compared with *in vivo* gene delivery where the whole foetus is

exposed to high titre gene vector constructs. Also, the manipulation of transduced cells outside of the receptor's body allows pre-transplantation monitoring, increasing the chances of transgene long-term gene expression. On the other hand, technical difficulties in obtaining enough stem cells together with laborious procedures of purification and contamination avoidance remain a challenge.

More recently the laboratory of Anna David reported a new approach using amniotic fluid stem cells as an alternative autologous source of cells with HSC potential. These can be readily harvested and the results obtained showed low foetal loss rates (Shaw *et al.,* 2011), but such haematopoietic potential has only been reported in the murine system and its still unknown whether a translation into human therapeutics would provide similar results.

1.5.5.2 Direct vector administration *in vivo*

In this approach the gene containing vector is transfected directly into the foetus, leading to *in vivo* transduction of HSCs in the liver.

Direct vector administration *in vivo* for the treatment of the haemoglobinopathies was first attempted in 1998 by IP retroviral delivery to early-gestation sheep foetuses. The authors reported long-term transduction of HSC (Porada *et al.,* 1998), and more recently by retroviral IP delivery to first trimester non-human primate foetuses (Tarantal *et al.,* 2006). This study reports that the vectors targeted mononuclear cells in the peripheral blood and bone marrow at a low frequency of between 0.5 and 4%.

To date the only report on *in utero* gene therapy for thalassaemia aims to correct a mouse model of α-thalassaemia (Han et al., 2007). In this study the authors deliver a lentiviral vector derived from the TNS9 vector (**Figure 1.13**). All elements in the original vector remained except for the replacement of the β-globin gene with either the human α-globin gene or a cDNA fragment containing a eGFP reporter.

Human α-globin gene expression was detected in the liver, spleen, and peripheral blood, reaching its peak at 3–4 months post-injection, reaching 20% in some recipients. However, the expression declined 7 months post injection.

Colony-forming assays performed on BM cells showed low abundance of the transduced human α-globin gene and the lack of its transcript, a demonstrating that the vector used could not transduce hematopoietic stem cells adequately to sustain gene expression. Regardless, this study proves that lentiviral vectors can be an effective vehicle for delivering the human α-globin gene into erythroid cells *in utero*.

1.5.6 Future challenges

Although great progress has been made there are many remaining challenges for pre-natal and cellular therapy. The proof-of-principle for foetal gene therapy for many disorders has been demonstrated in rodent and in large animals, but safety concerns regarding the risk of insertional mutagenesis means that no human *in utero* gene therapy trials have been reported or are planned at present (Coutelle, 2008). Further development of the technique regarding long-term efficacy as well as safety aspects, must be pursued in animal models before clinical trials in human foetuses can be initiated.

1.6 Aim of the Project

Gene therapy advances have been slow and disrupted by complications associated with transgene silencing, variable expression and PEV. Lentiviral vectors are promising tools for gene delivery due to their relatively large packaging capacity and ability to infect a range of different cell types. However, recent studies in the search for the optimal LVs once again raised safety concerns regarding insertional mutagenesis and potential future heath complications of patients submitted to gene therapy treatments (Cavazzana-Calvo *et al*., 2010). The search for the optimal vector makes more sense than ever, as the current lentiviral design for gene therapy for the haemoglobinopathies is still clearly far from ideal.

This project addresses the major issues associated with lentiviral gene therapy for the haemoglobinopathies and aims to develop a therapeutic globin vector capable of overcoming silencing and to provide sustained long-term, cell specific and reproducible expression with minimum genome disturbance.

1) Previous studies have successfully reported the use of ID-LVs to overcome integration disturbance in retina (Yanez-Munoz *et a.,* 2006), muscle (Apolonia *et al.*, 2009) and neurons (Rahim *et al.,* 2009). To date, only ubiquitous expressing transcription units have been studied and shown to provide good expression from ID-LV. We hypothesise if the efficient expression seen with ID-LV can be extended to a tissue specific therapy system. In particular, if ID-LVs are feasible to be used in an erythroid context.

2) *In utero* direct injection has been developed by others with encouraging outcomes in pre-clinical model systems addressing a broad spectrum of genetic disorders (Dejneka *et al.,* 2004: Rucker *et al.,* 2004: Waddington *et al.,* 2004) and one documented attempt at *in utero* gene therapy for α-thalassaemia (Han *et al*., 2007). We Hypothesize if sustained and reproducible level of expression can also obtained when an LV expressing human *HBB* is delivered *in utero* into wild-type mice, and if sufficient expression be obtained to attempt rescue a KO model of β-thalassaemia using and in utero approach.

Chapter Two

MATERIALS AND METHODS

2.1 Materials

2.1.1 Commercial kits

EndoFree Plasmid Maxi kit (QIAGEN®, *Hilden, Germany*)

HighSpeed Plasmid Maxi Kit (QIAGEN®)

Human Hemoglobin ELISA Kit (Bethyl Laboratoreis Inc., *Montgomery, TX USA*)

Luciferase Assay System (Promega, *Madison, WI, USA*)

QIAEX II Gel Extraction kit (QIAGEN®)

QIAquick Gel Extraction kit (QIAGEN®)

QIAprep Spin Miniprep kit (QIAGEN®)

QIAzol Lysis reagent (QIAGEN®)

QuickChange® II Site-Directed Mutagenesis kit (Stratagene® *Santa Clara, CA, USA)*

Quick Ligation ™ kit (New England Biolabs – NEB®, *New England, USA*)

RNeasy Mini kit (QIAGEN®)

TOPO® TA Cloning kit for Subcloning (Invitrogen™ , *Carisbad, CA, USA*)

Wizard® Genomic DNA purification kit (Promega)

2.1.2 Equipment

Avanti J-20 Centrifuge (Beckman Coulter, *High Wycombe, UK***)**

Becton Dickinson FACS Calibur (Franklin Lakes, *NJ, USA*)

BioDoc-It UV transilluminator system (Ultra-Violet products- UVP Ltd. , *Upland, CA, USA*)

Biofuge Pico microfuge (SORVALL ®, *New Castle, Delaware, USA)*

BioPhotometer (Eppendorf, *Hamburg, Germany*)

Boyle's apparatus (British Oxygen Company, *UK*)

Centrifuge Universal Legend RT (SORVALL ®)

Dyad thermocycler (MJ Research Inc., *St Bruno (Quebec), Canada*)

FACS Calibur (Beckton Dickinson and Company – BD, *Franklin Lakes, NJ, USA*)

Flourescence light Microscope, Eclipse TS100 (Nikon, *Melville, NY, USA*)

Gene Quant II spectrophotometer (GE Healthcare Bio-Science Corp., *Piscataway, USA*)

Gel Pulser/Micro pulser 0.2cm cuvettes (BioRad®, *Hercules, CA, USA*)

IVIS cooled charge-coupled device (CCCD) camera (Caliper Life Sciences, *Hopkinton, MA, USA*)

Legend RT centrifuge (SORVALL ®)

Light Microscope, TM (Nikon)

L8-60M Ultracentrifuge (Beckman Coulter)

Micro Centrifuge Pico Biofuge (SORVALL ®)

Multitron shaking incubator (Infors AG, *Bottmingen/Basel, Switzerland*)

Nano-Drop 1000D (Thermo Scientific, *Waltham, Massachusetts, USA*)

Protean Minigel system (BioRad®)

Research CO_2 Incubator (LEEC, *Silver Spring, Maryland, USA*)

Shaking Multitron Incubator (Infors AG)

Thermanox® coverslips (Nalge Nunc Int. *Penfield, New York, USA*)

Ultra-Turrax T25 homogeniser (Janke and Kunkel, IKA® *Labortechnik. Staufen, Germany*)

UV transilluminator (UVP Ltd.)

UV spectrophotometer Gene Quant II - Pharmacia Biotech Ultrospec 2000 (Analytical Instruments LLC, *Golden Valley, MN*)

2.1.3 General chemicals and reagents

2.1.3.1 General chemicals and reagents used in cloning procedures

Ampicillin (Sigma-Aldrich, *St.Louis, Missouri, USA*)

β-Agarase (Cambrex Corp. *East Rutherford, NJ, USA*)

Bovine serum albumin (Sigma-Aldrich)

Calf intestinal alkaline phosphatase - CIP (NEB®)

Chloroform (Sigma-Aldrich)

D-luciferin salt (Gold Biotechnology Inc., *Saint Louis, MI, USA*)

DEPC treated water (Invitrogen™)

DNA Rehydration Solution (Promega)

dNTPs, 2.5mM (Invitrogen™)

Ethanol (BDH Merck Ltd., *Poole Dorset, UK*)

Ethidium bromide (Sigma-Aldrich)

Glycerol (Sigma-Aldrich)

Isoflurane (Abbott Laboratories, *Abbot, IL, USA*)

Isopropanol (BDH Merck Ltd.)

$MgCl_2$, 25mM (Promega)

NEB buffer 1 (NEB®)

NEB buffer 2 (NEB®)

NEB buffer 3 (NEB®)

NEB buffer 4 (NEB®)

Phenol (Sigma-Aldrich)

Phenol : chloroform : isoamyl Alcohol 25 : 24 : 1, v/v (Invitrogen™)

Restriction endonucleases (NEB®)

SDS-running buffer (10x) (GeneFlow Ltd. *Staffordshire, UK*)

SYBR® Premix Ex Taq™ II mastermix (Lonza Group Ltd., *Basel, Switzerland*)

Taq Polymerase (Promega)

Tris Base (Sigma-Aldrich)

TE buffer solution (Invitrogen™)

Tris Base (Sigma-Aldrich)

Tris-HCl (BDH Merck Ltd.)

Tween-20 (Sigma-Aldrich)

T4 DNA ligase (NEB®)

T4 DNA ligase buffer (NEB®)

T4 polynucleotide kinase (NEB®)

T4 polymerase (NEB®)

2’ – deoxynucelotide triphosphates, 100mM (dNTPs) (GE Healthcare Bio-Science Corp., *Chalfont St. Giles, UK*)

2.1.3.2 Reagents used in tissue culture experiments

Dimethyl sulphoxide – DMSO (Sigma-Aldrich)

Foetal calf serum (PAA)

Dulbecco's Modified eagle Medium (DMEM) + Glutamax (PAA, *San Diego, CA, USA*)

Formaldahyde –PFA (Sigma-Aldrich)

Optimem® (Invitrogen™)

Phosphate buffered saline (PBS) (Sigma-Aldrich)

Phosphate buffered saline + Calcium and Magnesium (PBS +) (Sigma-Aldrich)

Polyethylenemine (PEI) (Sigma-Aldrich)

Polybrene (Sigma-Aldrich)

Magnesium cloride – $MgCl_2$ (Sigma-Aldrich)

Streptavidin/horse radish peroxidase (streptavidin-HRP) (DakoCytomation, *Glostrup Denmark*)

TMB+ substrate chromagen (DakoCytomation)

Tripsin TrypLE express (Invitrogen™)

Trypan blue (Sigma-Aldrich)

2.1.4 Plasmids vectors

pBluescript SK+ (Stratagene®)

pCR®2.1 TOPO® vector (Invitrogen™)

2.1.5 Bacterial strains

DH5α Chemically competent *E.coli* (Invitrogen™)

GeneHogs Electrocompetent *E.coli* (Invitrogen™)

TOP10 One Shot® Electrocompetent *E.coli* (Invitrogen™)

XL10 GOLD chemically competent *E.coli* (Stratagene®)

2.1.6 Prepared solutions

2.1.6.1. Solutions for cloning procedures

Loading buffer (DNA) (10x)

0.5 ml 1M Tris-HCL pH7.6 (BD)

25 ml glycerol (Sigma-Aldrich)

0.5 ml 10% sodium dodecyl sulphate - SDS (Sigma-Aldrich)

0.05 g Orange G (Sigma-Aldrich)

Make up to 50 ml with ddH_2O

LB (Luria-Bertani Media)

10 g bacto-tryptone (BD)

5 g bacto-yeast extract (BD)

10 g NaCl (BDH Merck Ltd.)

Make up to 1 Litre with ddH_2O

Sterilise by autoclaving

LB Agar

Add 15 g/ Litre bacto-agar (BD) to *LB*

Sterilise by autoclaving

50x Tris-Acetate-EDTA (TAE) 50x electrophoresis buffer

0.9ml Tris base (electrophoresis grade) (BD)

285.5 ml glacial acetic acid (BD)

500 ml 0.5M ethylenediaminetetraacetic acid (EDTA) (BD)

Make up to 50 L with ddH_20 and ajust to 7.2

Tris glycine buffer (10x)

14 g glycine (electrophoresis grade) (Sigma-Aldrich)

30 g Tris base (Sigma-Aldrich)

Make up to 1 Litre with ddH_20

200 ml methanol (BDH Merck Ltd.)

Make up to 1 Litre with ddH_20

Sodium acetate

3M sodium acetate adjusted to pH5.2 with glacial acetic acid

Sterilise by autoclaving

2.1.6.2 Solutions used for growing mammalian cells

Culture media	Description
Dulbecco's Modified Eagle Medium (DMEM)	DMEM supplemented with 4.5g L-ananyl-L-glutamine, 4.5µg/l of glucose, 10µg/ml of penicillin and streptomycin, 10% (v/v) heat inactivated foetal calf serum
1x phosphate buffered saline (PBS)	1 tablet of PBS (Sigma-Aldrich) at 7.5 dissolved per 100ml of ddH_2O

Table 2.1 - Solutions used in mammalian cell culture

2.1.6.3 Other solutions

Immunofluorescence buffer

PBS containing 0.1% sodium azide and 0.2% BSA

FACS buffer

PBS containing 4% paraformaldehyde

Lysis buffer (for animal procedures)

10mM Tris-HCl, pH7.5

10mM NaCl

3mM $MgCl_2$,

0.1% Triton X-100

0.1M sucrose

fresh 0.5mM dithiothreitol (DTT)

Protein extraction solution

PBS containing:

0.05% Tween20®

0.05% Triton X-100

SET buffer

50µl 10% sodium dodecyl sulphate - SDS

0.5mg/ml proteinase K

2.1.7 Enzymes

2.1.7.1. Restriction enzymes

Restriction enzyme	manufacture
Acc65 I	NEB®
Age I	NEB®
Avr II	NEB®
Afl II	NEB®
Bam HI	NEB®
Cla I	NEB®
Eco RI	NEB®
Eco RV	NEB®
Eco NI	NEB®
Hind III	NEB®
Hpa I	NEB®
Nco I	NEB®
Not I	NEB®
Pst I	NEB®
Pme I	NEB®
Sca I	NEB®
Sma I	NEB®
Sna BI	NEB®
Spe I	NEB®
Swa I	NEB®
Xho I	NEB®

Table 2.2 - Restriction enzymes used in this study

2.1.7.2 Other enzymes

Calf intestinal alkaline phosphatase – CIP (NEB®)

DNA Polymerase I - *E. coli* (NEB®)

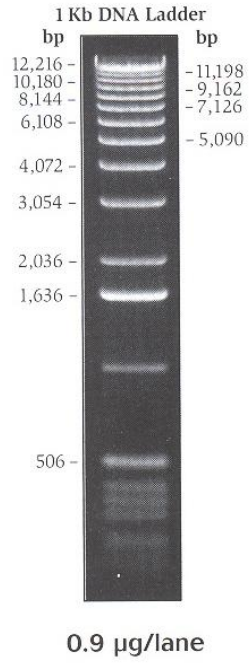
DNA Polymerase I, large (*Klenow*) fragment (NEB®)

DNase I - RNase-free (NEB®)

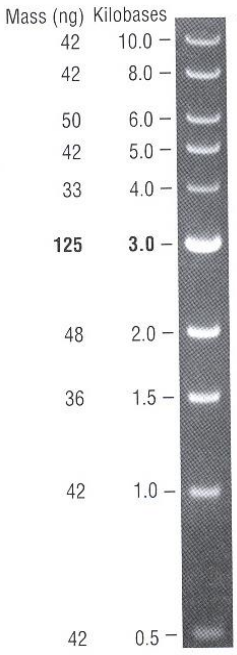
T4 DNA ligase (NEB®)

RNase I (NEB®)

2.1.8 DNA Molecular size markers



1kb molecular size ladder from Invitrogen™



1kb molecular size ladder from NEB®

Figure 2.1 - Molecular size markers

2.1.9 Tissue culture cell lines

Cell line / Species	**K562 (Human)**
reference	Lozzio and Lozzio, 1975
Derivation	myelogenous Leukemia cell line, which is bcr:abl positive Derived from a 53 year old female CML patient in blast crisis
Morphology	Leukaemia
Growth characteristic	Non-adherent (suspension cells)
Characteristics	Can spontaneously develop characteristics similar to early-stage erythorcytes, granukocytes and monocytes but easily killed by natural killer cells as they lack the MHC complex required to inhibit NK activity. Constitutive expression of episilon ε and γ globin

Cell line / Species	**Human Erythro-leukaemia (HEL)**
reference	Martin and Papayannopoulou, 1982
Derivation	Derived from the peripheral blood of a Hodgkin's disease patient.
Morphology	Leukaemia
Growth characteristic	Non-adherent (suspension cells)
Characteristics	Capable of globin synthesis and mainly expresses γ globin chains, small amount in embryonic chains (ζ & ε) but no β-globin chains. Although it does not express β globin, it provides an adequate erythroid environmenr for β globin transgene expression

Cell line / Species	**HEK 293T (Human)**
reference	Graham et al., 1977
Derivation	Derived from human embryonic kidney cells and transformed with adenovirus V DNA
Morphology	Fibroblast
Growth characteristic	Adherent
Characteristics	Contains a stably integrated SV40 Larger T-antigen gene, which allows for episomal replication of transfected plasmids containing the SV40 origin of replication.

Cell line / Species	**Murine erythro-leukaemia – MEL (Mouse)**
reference	Singer et al., 1974
Derivation	Pro-erythroid leukaemia cells transformed with Friend virus.
Morphology	Pro-proerythroblast APRT-tetraploid
Growth characteristic	Semi –adherent
Characteristics	dimethylsulphoxide (DMSO) to terminally differentiate along the erythroid pathway.

Cell line / Species	**C_2C_{12} (Mouse)**
reference	Yaffe & Saxel *et al.*, 1977
Derivation	myoblast cells from muscle
Morphology	fibroblast
Growth characteristic	Adherent
Characteristics	Can be induced to terminally differentiate into fused myofibres

Table 2.3 - Tissue culture cell lines used in this study

2.1.10 Antibodies

Antibody	Raised in	Method	Dilution	Company
CD3-PE	Rat	Cellular staining	1/500	BD Bioscience
CD11b-APC	Rat	Cellular staining	1/500	BD Bioscience
CD19-PerCP	Rat	Cellular staining	1/500	BD Bioscience
Ter119-APC	Rat	Cellular staining	1/500	BD Bioscience
Fc Block	Rat	Cellular staining	-	BD Bioscience
Anti-GFP antibody	Mouse	GFP-ELISA	1/10.000	AbCam, Cambridge UK
Biotin conjugated polyclonal anti-GFP	Goat	GFP-ELISA	1/5000	AbCam, Cambridge UK

Table 2.4 - Table with the antibodies used in this study

2.2.11 Computer programs and Internet pages

CellQuest™Pro software, *version 5.2.1*(BD)

Primer 3

SDS version 2.2.2© (Applied Biosystems, 2004)

Sequencing Analysis Software, *Version 5.1.1* (Applera)

National centre for Biotechnology Information (NCBI), http://www.ncbi.nlm.nih.gov

Vector NTI 10 (Invitrogen™)

7900HT Fast Real Time PCR System

Living Image software (Caliper Life Sciences)

2.2. Methods

2.2.1 Plasmid sub-cloning

In this context cloning refers to the generation of novel gene expression vectors and was accomplished as follows; appropriate insert and vector DNA fragments were covalently ligated. A ligation mixture containing the recombinant plasmid was transformed into bacteria and plated onto agar plates containing an appropriate antibiotic, where individual bacterial colonies grow in the presence of the circular plasmid replicon. Recombinant DNA was obtained through cell lysis and recovery of recombinant DNA with subsequent characterisation by restriction enzyme digestion and agarose gel electrophoresis. Large quantities of recombinant plasmid DNA were obtained by inoculating large cultures of the designed construct. Cell clones containing the correct plasmid were then stored long-term as a glycerol stock.

2.2.1.1 Preparation of plasmid DNA from bacterial hosts

Bacterial stocks of the DNA were streaked onto agar plates containing 100μg/ml of the appropriate antibiotic and grown overnight at 30 or 37ºC. Single colonies were subsequently picked and inoculated into 5 ml or 220-500 ml *Luria Bertani* broth (volume depends on the vector copy number) containing 100 μg/ml of the appropriate antibiotic for mini-preparation or maxi-preparation of DNA respectively. Cultures were expanded at 30 or 37ºC in a shaking incubator overnight. Mini- or maxi-preparation of DNA was performed using the appropriate QIAGEN® kit and eluted in ddH20 as per the manufacturers' protocols. Concentrations of DNA were then determined using a UV spectrophotometer as described in *section 2.2.1.8.*. Integrity of DNA was accessed by electrophoresis it out on a 0.8% agarose gel and stored at -20ºC.

2.2.1.2 DNA digestion by Restriction Endonucleases

Digests were performed as following;

Components	Volume
DNA	1-40µg DNA (concentration of DNA depended on individual experimental requirements)
appropriate Restriction Endonuclease (NEB®)	1 unit/µg DNA l
Buffer (1,2,3 or 4 NEB®)	10 of total reaction volumel
Nuclease-free Water	Todisered total volumel

Table 2.5 – Restriction Digests

Digestions with two enzymes were carried out simultaneously using a compatible buffer or performed sequentially with an intervening phenol-chloroform extraction and ethanol precipitation/purification step.

2.2.1.3 DNA blunting reactions

For generation of blunt-ended digested DNA products, T4 DNA polymerase was used. T4 DNA polymerase catalyses the synthesis of DNA in the 5' to 3' direction and also has 3' to 5'exonuclease activity. For blunt-ending of 5' overhang ends, 1 unit of T4 DNA polymerase was added per 1 µg of purified DNA, along with 100 mM dNTPs and 1x T4 ligase reaction buffer. The reaction was incubated at 12ºC for 15 m. To stop the reaction, a final concentration of 10 mM EDTA was added, which was then incubated at 75ºC for a further 15 min to heat inactivate the enzyme. For removal of 3' overhangs ends, the dNTPs were omitted from the reaction mix.

2.2.1.4 Isolation/Purification of DNA fragments

DNA fragments were separated by agarose gel electrophoresis (containing 0.005% ethidium bromide) and excised under long wavelength UV light (360 nm).

The QIAQuick Gel Extraction Kit (QIAGEN®) was used according to manufacturer's instructions. Briefly, the gel slice was solubilised and passed through an anion exchange chromatography column, followed by washing and elution of DNA from column. The DNA was then eluted using in ddH_2O.

2.2.1.5 Agarose gel electrophoresis

DNA was fractionated on 0.7% agarose gels made with 0.5x TAE gel buffer in the presence of ethidium bromide (0.5μg/ml final concentration). These were placed into an electrophoresis tank containing TAE buffer solution (refer to materials) with the same concentration of ethidium bromide as the gel.
Samples loaded were supplemented with gel loading dye (Orange G) and in case of small sample volumes, ddH_2O was added to a final volume of 20μl.

Routinely, at either side of each set of samples were loaded 1 kb DNA Ladder (*section 2.1.8*) used for sizing linear double-stranded DNA fragments by comparison with the migrating distances against the molecular size standard.

2.2.1.6 Phenol:Chloroform extraction

An equal volume of phenol : chloroform : isoamyl alcohol reagent was added to the DNA containing homogenate or solution and mixed thoroughly before centrifugation

for 10min at 13,000rpm. The upper, aqueous nucleic acid-containing phase was removed to a fresh tube without disturbing the proteinaceous layer at the interphase.

An equal amount of chloroform was then added to the decanted aqueous phase and mixed thoroughly before centrifugation for 10min at 13,000rpm to further remove residual phenol. The upper phase, containing purified DNA, was removed to a fresh tube without disturbing the chloroform lower phase. DNA was then purified further by ethanol precipitation step.

2.2.1.7 Ethanol precipitation

DNA solutions were made 0.3M with sodium acetate to which was then added 2.5 volumes of 100% ethanol. The solution was vortexed and incubated at -20 for up to one hour. Alternatively, 0.6 volumes of isopropanol were added to the nucleic acid solution and left to precipitate for 15 minutes.

Precipitated DNA was recovered by centrifugation and ethanol supernatant removed. The DNA pellet was then washed with 70% ethanol to remove isopropanol excess salt and again recovered by centrifugation. The pellet was left to air dry for 5 minutes before re-suspension in dd$H_2$0.

2.2.1.8 Estimation of nucleic acid concentration

The concentration of DNA samples was quantified by using a Gene Quant II spectrophotometer. An optical density (OD) reading at 260nm with an absorbance unit of 1.0 corresponds to approximately 50µg/ml for double stranded DNA and 40µg/ml for RNA.

Nucleic acid concentration was also measured using a Nano-Drop instrument. Volumes of 1µl of DNA / RNA were loaded for absorbance quantification.

Less accurate concentration estimates were obtained by electrophoresis on agarose gels and comparing the intensity of the desired DNA fragment against a commercially available and quantified DNA size ladder (section **2.1.8**).

2.2.1.9 Oligonucleotide hybridization

Complementary single stranded oligonucleotides obtained commercially were mixed in a 1:1 molar ratio in PNK buffer and denatured by incubation at 95ºC for 2min in ddH2O to a final volume of 50µl. The mixture was then allowed to cool down to 30ºC gradually to allow annealing of complementary strands. The reaction was then stored at -20C for later use.

2.2.1.10 Phosphatase treatment of vector

Ligation of DNA requires the formation of phosphodiester bonds between the 5'phosphate group and 3'hydroxyl group of the ends of the DNA strands. To prevent self-ligation of digested vector and the formation of vector-vector concatemers formed during ligation reactions when fragments have the same cohesive termini at both ends, the vector DNA should preferably be de-phosphorylated. Restriction digest reaction volumes were increased by 10µl with fresh 10x restriction endonuclease buffer, 10x BSA and dd$H_2$0. Calf-intestinal phosphatase (CIP) was then added at a concentration of 0.5 units per 1µg of DNA. For 5'overhang ends, CIP treatment was performed at 37ºC for 1 h. For 3'overhangs and blunt-ended DNA fragments, CIP-treatment was performed at 37ºC for 2 h. CIP cannot be fully heat-inactivated so CIP-treated DNA was either purified by phenol-chloroform extraction or following agarose gel electrophoresis.

2.2.1.11 TOPO® TA PCR cloning

Cloning of PCR products into the TOPO® TA cloning vector was performed as described in the manufacturer's protocol. PCR products were amplified using proofreading Expand High Fidelity^{PLUS} DNA polymerase. To incorporate suitable restriction sites into the PCR products for use in subsequent cloning procedures, restriction site recognition sequences were included into primer sequences.

2.2.1.12 Polyadenyladenylation of PCR products for TOPO® TA cloning

Since the high fidelity polymerase used to generate PCR products for TOPO® TA subcloning procedures does not adenylate the ends of the PCR products, adenylation of the PCR products was a necessary requirement for cloning into the TOPO vector.
The following mixture reaction was set in a sterile microfuge tube:

Components	Volume
Phosphorylated DNA Oligonucleotide	100 pmol (5 pmol/µl)
10X 5´ DNA Adenylation Reaction Buffer	2 µl
1 mM ATP	2 µl
Mth RNA Ligase	2 µl (100 pmol)
Nuclease-free Water	to 20 µl

Table 2.6 - Polyadenyladenylation of PCR products for TOPO® TA cloning

The adenylation reaction was left to Incubate at 65°C for 1 hour followed by enzyme inactivation by incubation at 85°C for 5 minutes.

2.2.1.13 Ligation of DNA fragments

The concentrations of insert and vector were determined by agarose gel electrophoresis.

Ligation of DNA fragments was performed using the Quick Ligation™ Kit according to manufacturer's instructions (NEB®). A total of 50 ng of vector were mixed with a 3-fold molar excess of insert. Total volume was adjusted to 10 µl with dH_2O and then to 20µl with 10 µl of 2X Quick Ligation Buffer. The reaction was briefly centrifuged after which 1µl of Quick T4 DNA Ligase were added to the reaction followed by, again, brief centrifugation.The reaction was left to incubate at room temperature (25°C) and transformed after 5 minutes.

Alternatively, T4 DNA ligase enzyme was used; For cohesive (sticky) ends, 1 µl of T4 DNA Ligase in a 20 µl reaction, and/or for blunt ends, 1 µl of T4 DNA Ligase in a 20 µl reaction for 2 hours. This reaction was left to incubate overnight at 16ºC.

2.2.1.14 Generation of electrocompetent *E.coli* cells

E.coli strain were streaked onto an agar plate and grown at 37ºC overnight. A single colony was then used to inoculate 10ml of LB and grown overnight at 37ºC, which was subsequently employed to inoculate 1L of 2x YT media the following morning. Cultures were grown to OD600 0.5-0.8 and then placed on ice for 15-30 min. Following centrifugation at 3700 xg for 15 min, pelleted cells were resuspended in 1L ddH_2O. Cultures were again centrifuged and the pellet resuspended in 500 ml of ddH_2O. After another spin, cell pellets were resuspended in 20 ml 10% glycerol. The cells were centrifuged once more then resuspended in 2 ml 10% glycerol and 40µl aliquots were made.

2.2.1.15 Transformation of bacterial competent cells by heat shock

A 50ng aliquot of DNA ligation product was mixed with 25µl of either TOP 10 or DH5α competent cells thawed on ice from their usual -80^{o}C storage, taking care to avoid repetitive pipeting and violent shaking. The mixture was left to incubate for 30 minutes followed by a heat shock at 37ºC for 60 seconds, to allow DNA uptake by cells. The sample was then returned on ice for a further 2 minutes after which 500µl of LB medium was added. The cells were allow to grow at 37ºC for 60 minutes in a shaking incubator (recovery phase) after which they were platted onto antibiotic containing agar plates and left to grow overnight at 37ºC. Colonies were picked with a sterile loop and used to inoculate 5ml of LB media supplemented with 100µg/ml antibiotic (ampicillin) followed by incubation at 37º C in a shaking incubator overnight.

2.2.1.16 Transformation of ligated DNA by electroporation

Competent cells (30-100ml aliquots) were thawed on ice for 10 minutes, to which was then added 2µl of the ligation products transfect to a pre-chilled electroporation cuvette and incubated on ice for a further 30-60s. Following placement in the Genepulser II apparatus, an electric pulse of 25µF capacitance, 2.5kV and 200 Ohm resistance was delivered to the cells. The cells were then placed in 950µl of LB broth and incubated for 15 minutes to 1 hour at 37ºC. Aliquots of 100µl of cells were then plated onto agar plates with ampicillin (100 mg/ml) and incubated overnight at 37ºC.

2.2.1.17 Storage of bacterial clones

A 500µl of cultured cell clones (known to contain the correct plasmid) was made with 16% glycerol and vortexed. This glycerol stock suspension was stored at -80^{o}C indefinitely.

2.2.1.18 Extraction of plasmid DNA by alkaline lysis

Solutions used for plasmid preparation:

Solution	Description	Function
P 1 (added 10μg/ml RNase A)	15mM Tris-HCl (pH 8) 10mM EDTA	bacterial pellet re-suspension
P 2	0.2M NaOH 1% (w/v) SDS	Lysis of bacterial membrane and denaturation of double-stranded DNA
P 3	3M potassium acetate (pH 5.5)	Neutralisation and renaturation of plasmid DNA (but not bacterial DNA)

Table 2.7 - Solutions used for extraction of plasmid DNA by alkaline lysis

Plasmid DNA was extracted from the bacterial cultures using the alkaline lysis protocol (Birnboim & Doly, 1979) as described below. From each bacterial culture 1.5ml were transferred into a microcentrifuge tube and spun for 1minute at 13 000rpm. Supernatant solution was removed and the cell pellet was resuspended in 100μl of P1 by repeated pipetting. P2 (100μl) was then added followed by gentle mixing by inversion and incubation at room temperature for 1-2 minutes, or until the mixture became clear (through bacterial lysis). P3 (100μl) was then added to the mixture, gently inverted until no viscous mixture was visible. The mixture was the subjected to a phenol:chloroform extraction followed by ethanol precipitation.

2.2.1.19 Large scale plasmid preparation

Plasmid DNA was prepared from overnight cultures grown in LB broth containing appropriate antibiotics (100μg/ml) that were inoculated previously with a

clonal population of bacteria. A large yield of purified plasmid DNA was obtained through use of column purification with the Plasmid Maxiprep Kit (QIAGEN®) according the manufacturers instructions. Plasmid constructs to be used for lentiviral vector generation, the Endofree plasmid Maxiprep kit (QIAGEN®) was used.

2.2.2. Tissue culture

2.2.2.1 Propagation of Adherent Cell Lines

All adherent cell lines were maintained in Dulbecco's modified eagle medium (DMEM) containing *GlutaMAX* supplemented with 10% (v/v) sterile, filtered fetal calf serum (*FCS*) and 1% (v/v) penicillin/streptomycin (referred to as *complete DMEM*). Cells were grown in T80 (80 cm^2) or T175 (175 cm^2) tissue culture flasks with vented caps in 37°C incubators in a 5% CO2 atmosphere.

Cells were passaged when 80-90% confluent (as indicated by orange indicator dye de-colourisation); Confluent adherent monolayers were detached from culture flasks by medium removal, washing with 1x PBS solution followed by incubation for 5 minutes at 37°C with trypsin/EDTA. Complete *DMEM* was added for resuspension and a proportion of the medium discarded depending on desired cell density. Fresh medium was added at the desired volume (depending on the flask used).

2.2.2.2 Propagation of Non-Adherent Cell Lines

Non-adherent cell lines were maintained in *DMEM* media containing *GlutaMAX* supplemented with 10% (v/v) FCS and 1% (v/v) penicillin/streptomycin (referred to as complete *RPMI*). Cells were grown in T25 (25 cm^2), T80 or T175 tissue culture flasks (standing upright) in 37ºC incubators in a 5% CO^2 atmosphere. Cells were passaged following media colour-change; the cells were transferred to 15 ml centrifuge tubes and centrifuged at 1200 rpm in a tabletop centrifuge for 5 minutes, washed with PBS, diluted 1:10 or 1:20 in fresh complete *DMEM* and transferred to new tissue culture flasks.

For confluent suspension cells, medium was directly discarded and replaced with fresh medium.

2.2.2.3 Freezing of tissue culture cells in Liquid Nitrogen

A cell pellet containing was obtained as previously described. This pellet was re-suspended in 1ml growth medium containing 10% DMSO and transferred to a 1.5 freezing vial. This vial was then placed on dry ice for 2 hours followed by transfer to liquid nitrogen

2.2.2.4 Recovery of cells from Liquid Nitrogen storage

Vials were thawed quickly at 37ºC from liquid nitrogen storage and diluted in 10ml of complete medium. The cell suspension was then centrifuged at 1500 rpm for 5 minutes with subsequent removal of DMSO containing supernatant and re-suspension of the cell pellet in growth medium, and cultured as before

2.2.2.5 Cell counting

A cell suspension in 10ml of complete medium was prepared and cells counted with a standard haemocytometer.

2.2.2.6 Stable Transfection of Mammalian Cells by Electroporation

Plasmid DNA for transfection was isolated using the QIAGEN® Plasmid Maxi Extraction Kit. The plasmid DNA was linearised using a suitable restriction endonuclease. Phenol-chloroform extraction and ethanol precipitation was then carried out and the DNA pellet dissolved in 100μl of electroporation buffer (*DMEM* without additives).

Transfection was carried out by electroporation (Chu *et al.*, 1987) with 1-3x10^7 cells in log phase growth. The cells were harvested as described above and resuspended in 700µl of additive-free *DMEM*. This was added to the 100µl DNA mixture and incubated at room temperature with periodic shaking for 10 minutes. The DNA was introduced into the cells by electroporation using the BioRad Gene Pulser Unit, set to deliver a transient electrical pulse of 975µF at 250V. After electroporation, the cells were allowed to stand at room temperature for further 10 minutes before transfer to fresh complete *DMEM* and incubation at 37°C.

Stably transfected cells were selected for and maintained by growth in the presence of an appropriate antibiotic (geneticin sulphate) at a concentration of 800µg/ml, added 24 hours post transfection.

2.2.2.7 Induction of differentiation of MEL cells

Murine erythroleukaemia (MEL) cells were induced to undergo terminal erythroid differentiation by maintaining a log phase growth culture in complete DMEM containing 2% DMSO (v/v) for four days.

2.2.3 HIV-1 based Lentiviral vector production

Production and assays of second generation lentiviral vectors was conducted following Naldini protocol (1996). Lentiviral particles were generated by transient transfection of HEK293T cells seeded in T175 flasks so that the cells were 80-90% confluent at the time of transfection ($2x10^7$ cells per flask).

The following amounts of DNA were added to 5 ml of Optimem per T175 flask and filtered through a 0.2 µm acetate membrane syringe filter (Nalgene).

Plasmid vector	**Quantity**
Transfer vector	50µg
PMD2.VSVG (VSV-G envelope)	16µg
Packaging plasmid	32µg

Table 2.8 – Plasmids used in lentiviral vector production

1µl of a 10 mM stock of the synthetic polycation transfection reagent polyethylenimine (PEI) was added to 5ml of *Optimem* and filtered through a 0.2µm filter. DNA and PEI solutions were then mixed in a 1:1 ratio and incubated at room temperature for 20 min. The cells were washed with 10 ml *Optimem* and 10 ml of the PEI/DNA complexes were then added per T175 flask. The cells were incubated at 37ºC for 4hours followed by replacement of the transfection medium with 25 ml of complete DMEM media. The flasks are then left to incubate overnight.

2.2.3.1 Harvesting of lentiviral particles

Around 48 hours post-transfection the medium was harvested and replaced with 20ml fresh stock so the virus could be harvested for a second time 72h post-transfection. The harvested medium was cleared of debris by low-speed centrifugation at 2500rpm for 10 minutes and filtered through a 0.2 µm filter and then centrifuged at 4ºC for 24h at 3700xg. The supernatant was decanted and the tubes were inverted onto tissue to drain the remaining supernatant. 50 µl of *DMEM* without supplements or *Optimem* was added to the viral pellet and incubated on ice for 20 min. The viral pellet was then resuspended by repeated pipetting and centrifuged for 10 minutes at 4000rpm to remove cell debris and impurities,. 5-10 µl aliquots were then stored at -80ºC.

2.2.3.2 Lentiviral vector titration

For each viral harvest, 4 wells of a 24-well plate were seeded with $2x10^5$ HEK293T cells and transduced with serial dilutions of concentrated virus in order to achieve a range of multiplicity of infections (MOI) between 1.0 and $1x10^{-3}$, assuming that the initial viral titre was $1x10^8$ infectious units (IU)/ml. The MOI indicates the number of viral particles per cell. Specifically, 2µl viral concentrate was added to the first well to achieve an assumed MOI of 1, with a concentration of $2x10^5$ IU per ml. For the remaining wells, 1/10 serial dilutions of this well were carried out to achieve an assumed MOI of $1x10^{-1}$, $1x10^{-2}$ and $1x10^{-3}$ respectively. To the 1 ml in each well, a further 1 ml of DMEM medium was added containing 16µg/ml (w/v) of the cationic polymer polybrene to give a final concentration of 8µg/ml. The 5th well was left untransduced to act as a negative control.

At 72h post-transduction, the cells were harvested (using trypsin if adherent cells were used for titration) and centrifuged at 1000rpm for 2 minutes. The cell pellet was resuspended in 800µl 1% formaldehyde in PBS and transferred to an appropriate

fluorescence activated cell sorting (FACS) analysis tube. Tubes were covered in foil and stored at 4°C to awaiting FACS analysis

2.2.3.2 Image capture of eGFP reporter gene expression

Cells expressing enhanced green fluorescence protein (eGFP) were examined under a Nikon Eclipse TS100 fluorescence microscope via blue light excitation using a Osram 50W mercury vapour bulb. Fluorescent images were taken using an Olympus SP-350 8MG pixel digital camera.

2.2.3.4 Viral Titer quantification

Viral titre was simultaneously calculated by both FACS and from integrated viral vector copy number in tissue/cell genomic DNA by quantitative polymerase chain reaction (qPCR).

2.2.3.4.1 Quantification of viral titre by Fluorescence Activated Cell Sorting (FACS) analysis

Individual samples were analysed with a Beckton Dickinson FACS Calibur instrument and CellQuest™Pro software. Levels of eGFP fluorescence were analysed by detecting light emission at 525nm wavelength (colour spectrum range FL1), using untransduced cells to gate for positive signal. Specifically, samples of at least 10^4 cells suspended in 800µl 1% formaldehyde in PBS, were analysed for various parameters followed by construction of plots to visualise the analysis. Using parameters of size versus granularity, untransduced cells were used to gate particular subsections of the cell population such that only healthy, live cells were included in the analysis. Similarly, using parameters of light emission at 525nm (FL1) versus emission at 575nm (FL2),

untransduced cells were used as a negative reference to gate for positive eGFP signal. Settings, such as voltage and current, were then fine-tuned according to a known eGFP positive cell sample to avoid overlap in the FL1 (green fluorescence) and FL2 (red fluorescence) wavelength ranges, as true eGFP florescence is determined by emission at 525 nm (FL1). As untransduced cells autofluoresce to some extent in the FL2 range, it is important that detection of each range remains independent of one another to ensure accurate detection of eGFP fluorescence. Therefore, detection in the FL1 range was adjusted manually so that it was independent of emission in the FL2 range.

For the assessment of viral titre, a cell sample transduced at a predicted MOI $1x10^{-1}$ would be expected to show a 10% eGFP positivity level. Similarly, a sample of predicted MOI $1x10^{-2}$ would be expected to have a positivity level of 1% and so on. Any deviation from this expected percentage positivity would thus be used to determine the actual titre. For example, if the sample of predicted MOI $1x10^{-2}$ has a 2% positivity level, the titre in this case would in fact be $2x10^{8}$ IU per ml, rather than the estimated $1x10^{8}$ IU per ml. Viral titre was calculated from the sample/s showing between 5-10% GFP positive cells (MOI 0.05-0.1). This is because in a cell sample of low transduction giving 5-10% cells, signal from each GFP positive cell is more likely to derive from a single LV integration event rather than from multiple copies. A 5-10% positivity of a cell sample is therefore more likely to be directly proportional to viral copy number (that is, a 1:1 ratio). This ensures that a more accurate titre can be calculated.

2.2.3.4.2 Quantification of viral titre by TaqMan qPCR

As an alternative to titre quantification by FACS, titre can also be quantified by determining integrated viral vector copy number in tissue/cell genomic DNA or un-integrated proviral vectors by real-time quantitative polymerase chain reaction (qPCR; TaqMan). This methodology for titer estimation is particularly important for therapeutic vectors, in which the vector promoter does not drive a reporter gene but a therapeutic

cassette. Total genomic DNA was extracted from transduced tissues or cells and then diluted to an optimal concentration (20-50ng/µl).

TaqMan probes depend on the 5'- nuclease activity of the DNA polymerase used for PCR to hydrolyse an oligonucleotide that is hybridised to the target amplicon. TaqMan probes are oligonucleotides that have a fluorescent reporter dye attached to the 5' end and a quencher coupled to the 3' end. Probes are designed to hybridise to an internal region of the PCR product. In the unhybridised state, the proximity of the fluor and the quench molecules prevents the detection of fluorescent signal from the probe. However, during the PCR reaction, when the polymerase replicates a template on which a TaqMan probe is bound, the 5'- nuclease activity of the polymerase cleaves the probe, which decouples the fluorescent and quenching dyes (refer to **Figure 2.1,** below) Fluorescence naturally increases in each cycle, proportional to the amount of cleavage probe.

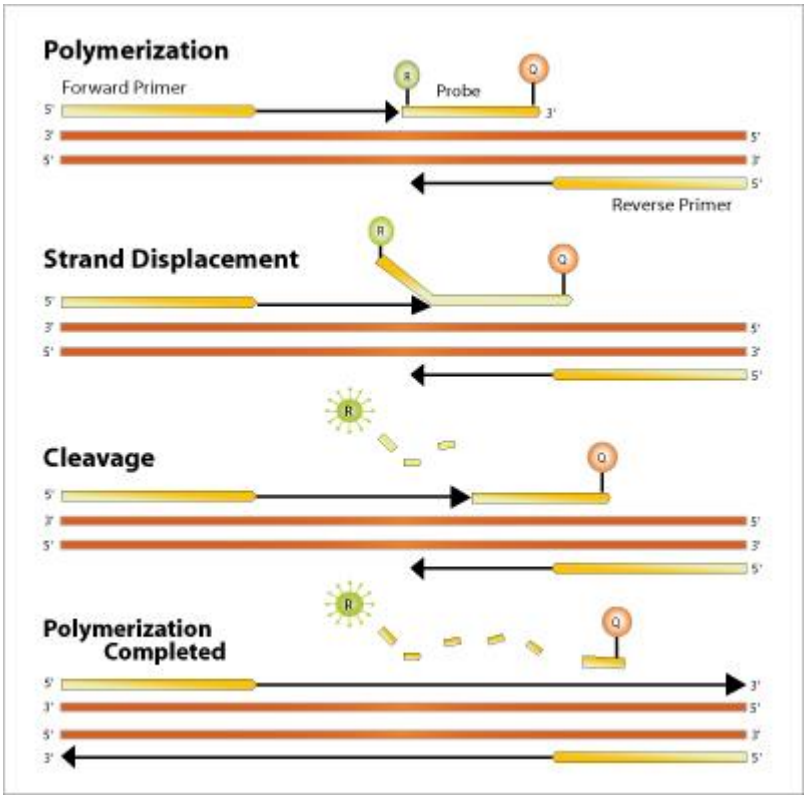


Figure 2.2 - Taqman PCR assay

(image taken from http://www.servicexs.com/plaatjes/TaqMan_RT-PCR_assay)

The qPCR reactions were set-up in a 96-well reaction plate and analysed by using Bushman's late reverse transcriptase (LRT) primer/probe set (Bushman et al., 2001); Bushman primers were designed to amplify a 120bp fragment of the HIV vector backbone between the right 5' LTR sequence and the 5' end of the *gag* gene, downstream of the *cppt* element, such that only integrated DNA or provirus forms that have completed the two template switches of reverse transcription are detected (in **Figure 2.2**, below) these were designated in the host lab as the "LRT primers"

As a reference, a clonal cell line with an independently determined viral copy number of one was used as a standard and repeatedly loaded together with samples analysed to act as internal reference. Both standards and samples were diluted to concentrations between 20-40ng/µl in deionised distilled water (ddH_2O) – these were accurately determined using the *Nanodrop* Sofware, although variations in the samples concentrations are not likely to affect the accuracy of the analysis as the reaction was performed in parallel with an endogenous internal control.

Primers and probes used:

LRT forward: 5' TGTGTGCCCGTCTGTTGTGT 3'

LRT reverse: 5' GAGTCCTGCGTCGAGAGAGC 3'

LRT probe: 5' – (FAM) – CAGTGGCGCCCGAACAGGGA – (TAMRA) 3'

Volumes loaded per well:

Total volume: 20µl

DNA sample: 2µl

Forward primer: 0.375 µl (3.75 µM of a 10 µM/µl stock)

Reverse primer: 0.375 µl (3.75 µM of a 10 µM/µl stock)

LRT probe: 1µl (2mM of a 2 µM/µl stock)

TaqMan Master Mix with ROX (PrimerDesign): 10 µl

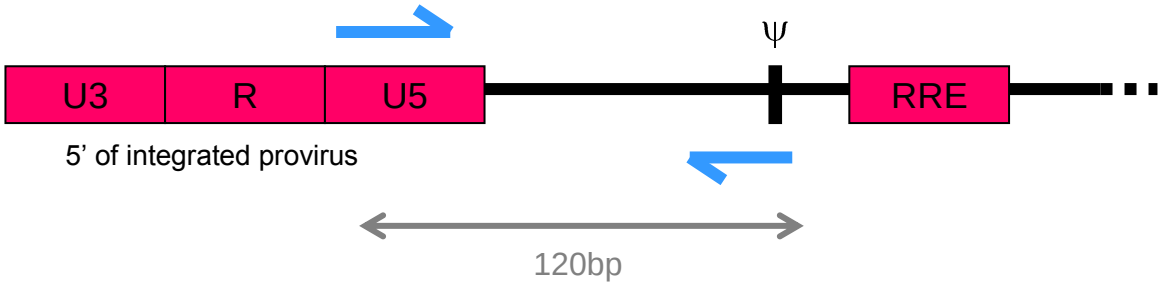


Figure 2.3 - Diagram showing position of Bushman's (LRT) qPCR primer/probe set

The endogenous human (*GAPDH*) or mouse (*Gapdh*) glyceraldehyde-3-phosphate dehydrogenase gene was used to determine absolute numbers of genomes present in each qPCR reaction using a primer/probe set amplifying an 73bp fragment of the 5' end of exon V of the exogenous GAPDH gene (refer to **Figure 2.3** below).

GAPDH Primers and probes used:

GAPDH forward: 5' ACCACAGTCCATGCCATCACT 3'

GAPDH (human) reverse: 5' GGCCATCACGCCACAGTT 3'

GAPDH (mouse) reverse: 5' GGCCATCACGCCACAGCTT 3'

GAPDH probe: 5' – (FAM) – CCACCCAGAAGACTGTGGATGGCC – (TAMRA) 3'

Volumes loaded per well:

Total volume: 20µl

DNA sample: 2µl

Forward primer : 1 µl (3.75 µM of a 10 µM/µl stock)

Reverse primers: 1 µl (3.75 µM of a 10 µM/µl stock)

GAPDH probe: 2µl (4mM of a 2 µM/µl stock)

TaqMan Master Mix with ROX (PrimerDesign): 10 µl

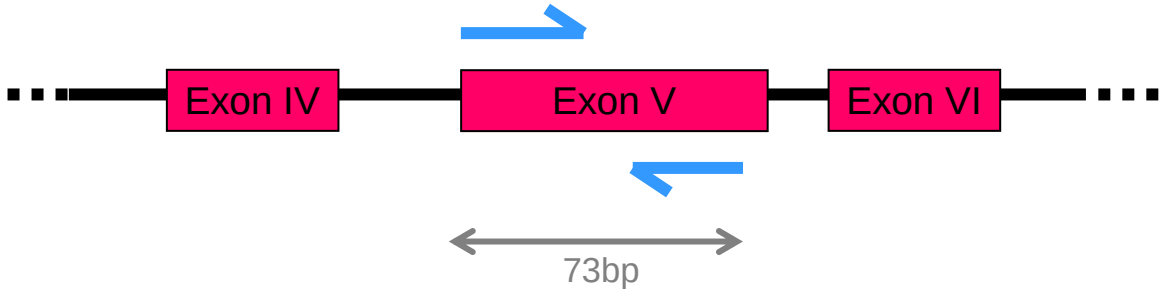


Figure 2.4 - Diagram showing position of the Glyceraldehyde-3-phosphatase dehydrogenase (GAPDH) qPCR primer set

The PCR reaction was conducted under the following cycling conditions:

Stage 1: 50℃ 2min

Stage 2: 95℃ 10min

Stage 3: 95℃ 15sec and 60℃ 1min (40 cycles)

Quantification was carried out by Sequence Detection Systems software (SDS version 2.2.2©) by reference to the standard curves of the reference with a known viral copy number.

2.2.3.5 Lentiviral vector cell line analysis

All lentiviral vectors were tested *in vitro* on appropriate cell lines, erythroid specific, pro-erythroblast MEL cells, and non-erythroid specific cell-line, HEK293T cells. Cells were seeded at a density of 5 x10^5 cells per well 12h prior to transduction at MOI of 1 or less depending on the vectors experimental requirements.. Transduction was carried out in 12 wells plates with total volume of 1ml medium. Cells were harvested

72h post-transduction for genomic DNA extraction (**2.2.5.1.1**), RNA extraction (**2.2.5.1.1**), total protein extraction (**2.2.5.1.1**) and FACS analysis (**2.2.3.4.1**). Ongoing cell line analysis and erythroid induction were collected at different time points (eg. day 3, 9 and 21 post-transduction).

2.2.4 Animal experiments

All animal work was carried out under United Kingdom Home Office regulations and was complaint with the guidelines of the Imperial College London ethical review committee. All animal experiments were carried out according to the institutional guidelines for animal use and manipulation. .

2.2.4.1 Neonatal injections

Neonatal delivery was carried out on wild type MF1 strain mice. Gene transfer was performed on hypothermia anaesthetised neonatal mice by intra-vascular administration through the superficial temporal vein within the first 48 h of life using a 32-gauge needle and transdermal approach (Chandler & Venditti, 2009). A volume of 40ml of viral suspension with 1-5x109 IP per ml were injected per pup.

Mice were marked for identification either by tail clipping or foot carbon marking and allowed to recover in a warm cage. Injected mice were bled via tail vein post-injection and were sacrificed at appropriate time-points when individual tissues were collected for molecular analysis. All mice showed normal tissue morphology.

2.2.4.2 *In utero* injections

In utero delivery was carried out on wild type MF1 strain mice. Pregnant females at the appropriate desired gestational age were anesthetised with Isoflurane. A midline laparotomy (1-1.5cm) was performed to expose horns of the gravid uteri. The yolk sac vessels of individual embryos were visualised under a dissecting microscope. Vector injection was done by inserting a glass needle (70-80µm) attaching to an Hamilton microliter syringe. A volume of 20ul of 1-5x10^9 infectious viral particles was injected

into each embryo via the vitteline vessel. Injected foetus were marked *in utero*, for post-partum identification, by injection of colloidal carbon into the flank.
After injection, the uteri were returned into the abdominal cavity and the abdomen was closed with silk suture. Mice were allowed to recover in a warm cage.

Born pups were marked for identification after birth either by tail clipping or foot carbon marking. Injected mice were bled via tail vein post-injection and culled at appropriate time points for molecular analysis

2.2.4.3 *In vivo* Blood collection

Animals were placed under isoflurane anesthesia, and with a scalpel a small excision was cut on the base of the tail. Blood samples were added to anticoagulant solution (0.109 M sodium citrate buffer) vortexed and centrifuged. Samples were kept on ice until analysis

2.2.4.4 Animal sacrifice by cardiac puncture

Mice were sacrificed by cardiac puncture so that the maximum amount of circulation blood was obtained, 250days after *in utero* injection. Briefly animals were placed under isoflurane anaesthesia and a 22-gauge needle coupled with a 1ml syringe was inserted in the centre of the mouse thorax. Blood was withdrawn slowly to prevent the heart collapsing.
As before, blood samples were added to anticoagulant solution (0.109 M sodium citrate buffer) vortexed and centrifuged; and samples were kept on ice until analysis.

2.2.4.6 Homogenate of tissue cells

Individual tissues were embedded in PBS solution and macerated using the sonicator to obtain a smooth solution, without suspension particles, from which DNA could be extracted

2.2.4.7 Extraction of Bone marrow cells

Bones were carefully cleaned from adherent soft tissues and the tip of each bone was removed with a scalpel. The marrow was harvested by inserting a 27-gauge syringe needle into one end of the bone and flushing with DMEM solution.
Individual cells were embedded in PBS solution from which DNA could be extracted

2.2.5 Gene Expression analysis

2.2.5.1 Sample extraction from Tissue culture cells and mouse tissues

2.2.5.1.1 Extraction of Genomic DNA from Mammalian Cells

Approximately $1x10^7$ cells were harvested and resuspended in 1 ml SET buffer and incubated at 42°C for 16h for protein denaturation. Genomic DNA was then recovered by phenol: chloroform extraction and ethanol precipitation. Alternatively the Promega Wizard DNA Purification Kit was used to extract genomic DNA from the cell pellet according to manufacturer's instructions.

2.2.5.1.2 Extraction of RNA from mammalian cells

Total RNA was extracted from a minimum of 5 $x10^6$ cells, harvested and resuspended in 500µl of TRI Reagent in accordance with the manufacturer's instructions.

2.2.5.2.3 Extraction of Protein from Mammalian Cells

Cells (10^6-10^7) were resuspended in protein extraction solution and incubated for 10 minutes at room temperature. Tissue (1-50mg) was homogenised in 500µl protein extraction solution. The cell/tissue homogenate was then centrifuged for 5 minutes at 13000rpm. The supernatant was removed and stored at -80°C for protein analysis.

2.2.5.2 Vector copy number determination

Analysis of samples for vector copy number determination was performed as described previously for titer estimation, using LRT and GAPDH primers (section 2.2.3.4.2). Copy number was determined by comparison with a reference sample known to contain a single viral copy per genome.

2.2.5.3 Reverse Transcriptase Polymerase Chain Reaction (RT-qPCR)

2.2.5.3.1 cDNA synthesis

cDNA was synthesised from total RNA using SuperScript® III Reverse Transcriptase and random hexamers as per the manufacturer's instructions. Briefly, cDNA synthesis reactions contained 1 μg of total RNA with 100ng of random hexamers and 100 mM DTT which was incubated at 94°C for 90 sec. Following incubation on ice for 2 min, reverse transcription was carried out in a final volume of 20 μl in the presence of 1mM dNTPs, 10 URNAsin, 200 U molony murine leukemia virus reverse transcriptase and 1st strand buffer.

Cycling conditions were

23°C 10 min

37°C 40 min

and 94°C 5 min

Following synthesis samples were diluted 1:6 in sterile water and stored at -20°C until required.

2.2.5.3.2 Reverse transcriptase PCR

Reverse transcriptase PCR reactions were carried out in a final volume of 20 μl containing 3 μl cDNA, 10 μl Thermostart PCR Sybergreen (SYGR) mastermix

(containing 2.5mM $MgCl_2$) and 2 µl DMSO. SYBR Green binds double-stranded DNA, and upon excitation emits light. Thus, as a PCR product accumulates, fluorescence increases.

The designed primers span a 164bp region starting from the 3' end of Exon2 of the *HBB* gene to exon 3, allowing amplification to occur only in cDNA fragments (**Figure 2.4**)

HBB RT-PCR primers:

RT-PCR globin gene Forward: GATCCTGAGAACTTCAGGCT (20)

RT-PCR globin gene Reverse: AGGCAGAATCCAGATGCTCA (20)

Thermocycling conditions

95ºC 15 min,

x25 95ºC 10 s

65ºC 40 s

72ºC 40 s

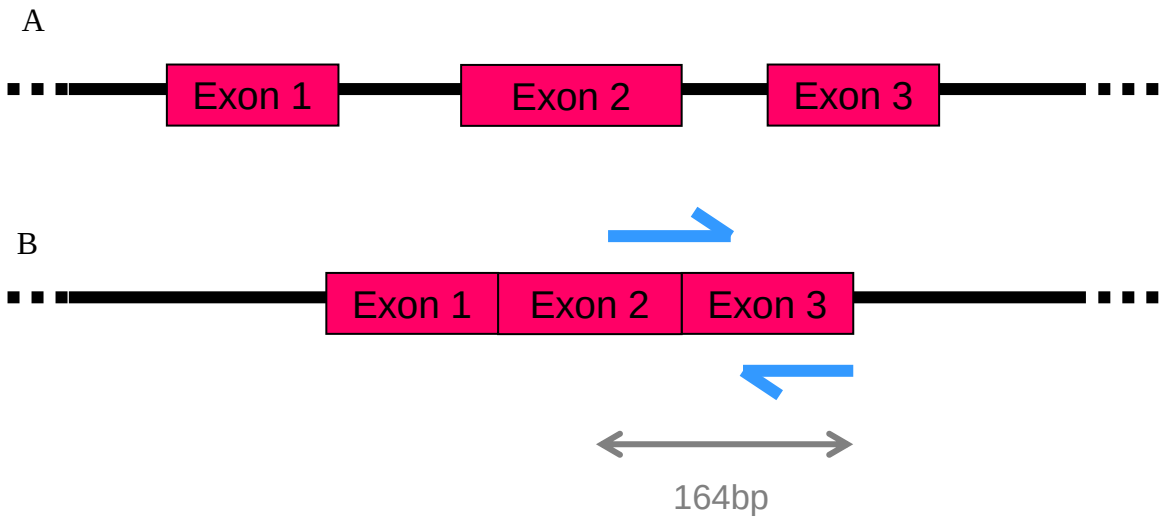


Figure 2.5 - Diagram showing position of the *HBB* RT-qPCR primer set. The primers are designed to span a 164bp region starting from the 3' end of Exon2 of the *HBB* gene to exon 3. **A** *HBB* DNA, not amplyifeid by the set of primers **B** cDNA fragment that can be amplified by the primer set.

GAPDH internal primers used are described in section **2.2.3.4.2, Figure 2.3.**

2.2.5.3.3 Expression levels quantification

Quantification was carried out by Sequence Detection Systems software (SDS version 2.2.2 ©Applied Biosystems, 2004). Taqman amplification results in amplification curves of which log phase depended on initial amounts of vector in the sample. The estimation of expression levels was obtained by comparison with standard curves obtained with reference sample (induced sample) of known concentration.

2.2.5.4 Sandwich Enzyme linked Immunosorbent Assay (Sandwich ELISA) for eGFP Quantification

eGFP ELISA was carried out by using Matched-Pair Antibody Set for ELISA. Briefly, eGFP protein was captured by the coated eGFP specific affinity-purified primary antibody on the wells of a microtitre plate. Captured proteins bound to a peroxidase conjugated secondary antibody expressed by incubation with TMB+ substrate chromagen. Upon development time points of 10 and 20min the green colour indicative of protein presence was quantified with a microplate reader at the wavelength of 490nm and compared with standards. The colour generated is proportional to the concentration of GFP present in the sample. Each sample was duplicated in the assay.

Wash buffer	0.05% Tween20® in 1XPBS
Bicarbonate buffer	0.16% Na_2CO_3 (w/v), 0.29% $NaHCO_3$ (w/v)
Block buffer	1% BSA (w/v) in 1XPBS

Table 2.9 - Constituents of Buffers Used in eGFP ELISA

2.2.5.4.1 Preparing Protein Samples

Total protein from both cells and tissue transduced with eGFP lentiviral vectors was quantified using the colorimetric Bio-Rad Protein Assay Kit I, which uses the Bradford Assay for protein quantification (Bradford *et al.*, 1976). Absorbance was measured at 595nm and the quantification of protein levels per sample was determined by calibration of standard curves created with known standard references of Bovine

serum albumin (BSA). Standard references were used to create a standard curve which was used to fit the absorbence data for known concentrations of BSA to a linear regression model using Microsoft® Excel (where $R^2 \geq 0.9$).

Calculation of protein concentration of unknown samples was then carried out. Aliquots of 2µg were then prepared for eGFP ELISA by addition of wash buffer to a total volume of 100ul.

2.2.5.4.2 Application of Primary Antibody

Mouse monoclonal anti-eGFP antibody was diluted 1/10,000 in bicarbonate buffer and 100µl added to each well of a flat bottomed, high binding 96 well plate, which was sealed and incubated overnight at 4°C, followed by washing with 250µl of wash buffer (repeated three times). Block solution (300µl) was then added to each well and the plate incubated for a minimum of 1h at 37°C followed by washing with 250µl of wash buffer (repeated three times). Then, 100µl of sample or standard in wash buffer was added to each well, including a negative control of wash buffer alone.

2.2.5.4.3 Preparing GFP Standard Curve

1µg of recombinant eGFP Protein was diluted to 1ml wash buffer. Then, 40µl of this was added to 1 ml wash buffer and serial dilutions carried out to create the following range of GFP dilutions; 40, 20, 10, 5, 2.5, 1.25, 0.63, 0.32 pg/µl to act as standard references.

After addition of 100µl of sample and standards (in triplicate), the plate was then sealed and incubated for 1h at 37°C followed by washing with 250µl of wash buffer (repeated three times).

2.2.5.4.4 Application of Secondary Antibody

Biotin-conjugated goat polyclonal anti-eGFP secondary antibody (Abcam®) was diluted 1/5,000 in block solution and 100µl was added to each well. The plate was then sealed and incubated for 1h at 37°C followed by washing with 250µl of wash buffer (repeated three times). Streptavidin/horse radish peroxidase was then diluted 1/5,000 in block solution and 100µl added to each well, containing the primary antibody/protein sample/secondary antibody 'sandwich'. The plate was then sealed and incubated for 1 hour at 37°C followed by washing with 250µl of wash buffer (repeated three times). Then 100µl of TMB+ substrate chromagen was added to each well and the plate incubated for 20-30 minutes at room temperature in the dark to allow colour change of the biotin-bound substrate chromagen catalysed by HRP. A 100µl volume of 3M H_2SO_4 was then added to each well and the plate gently tapped to ensure thorough mixing. Addition of H_2SO_4 ends the HRP-catalysed colour change, the extent of which depends on eGFP levels. Absorbence of light of 450nm wavelength was then measured for each sample on the Bio-Rad model 680 Microplate Reader at 10 and 30 minutes timepoints.

2.2.5.4.5 Fitting the Data to a Linear Regression

The eGFP quantification of unknown protein samples was carried out by fitting the absorbance measurements of eGFP protein standard to a linear regression using Microsoft® Excel. This standard curve (where $R^2 \geq 0.9$) was then used to calculate eGFP concentrations of unknown samples per mg of total protein.

2.2.5.5 Sandwich Haemoglobin ELISA for *HBB* quantification

Haemoglobin ELISA was carried out by using Human Hemoglobin ELISA Kit . Briefly, this kit is based on a sandwich ELISA; Human hemoglobin present in the test sample is captured by anti-human hemoglobin antibody that has been preadsorbed on the surface of microtiter wells. After sample binding, unbound proteins and molecules are washed off, and a biotinylated detection antibody is added to the wells to bind to the captured hemoglobin. A strepavidin-conjugated horseradish peroxidase (SA-HRP) is then added to catalyze a colorimetric reaction with the chromogenic substrate TMB (3,3',5,5'-tetramethylbenzidine). The colorimetric reaction produces a blue product, which absorbance at 450 nm is proportional to the amount of hemoglobin analyte present in the sample and a four-parameter standard curve can be generated. The hemoglobin concentrations in the test samples can then be quantified by interpolating their absorbance from the standard curve generated in parallel with the samples.

2.2.5.5 1 Sample dilution and standards preparation

Whole Blood was diluted as recommended per manufactur's instructions (1:50,000). For standards preparation 200-ng human Haemoglobin was reconstitute with 1.0 ml of 300 µl of 1X Dilution Buffer and serial dilutions carried out to create the following range of human haemoglobin 66.6, 22.2, 7.4, 2.5, 0.82, 0.27, and 0 ng/ml to act as standard references.

After addition of 100µl of sample and standards (in triplicate), the plate was then sealed and incubated for 1h at 37°C followed by washing with 250µl of wash buffer (repeated three times).

2.2.5.5 2 Incubation with Detection Antibody

A volume of 100µl of Detection Antibody to each well containing standard, sample or blank. The well plate was sealed and incubated for 1 hour at room temperature, 20-25°C, followed by washing with 250µl of wash buffer (repeated four times).

2.2.5.5.3 Incubation with strepavidin-conjugated horseradish peroxidase HRP Solution

A volume of 100 µl of HRP Solution was added to each well containing standard, sample or blank, which was sealed and left to incubate plate for 30 minutes at RT, followed by washing with 250µl of wash buffer (repeated four times).

2.2.5.5.4 Incubation with chromogenic substrate (TMB) and Stopping the Reaction

A volume of 100µl of TMB Substrate Solution into each well and the enzymatic reaction was allowed to develop a blue color at RT in the dark for 30 minutes. The reaction was stopped by adding 100µl of Stop Solution to each well.

2.2.5.5.5 Fitting the Data to a Linear Regression

HHB quantification of unknown protein samples was carried out by fitting the absorbance measurements of Haemoglobin protein standard to a linear regression using Microsoft® Excel. This standard curve was then used to calculate Haemoglobin concentrations of unknown samples per ng of total protein.

2.2.5. 6 Bioluminescence detection assays

2.2.5.6.1 *In vivo* Bioluminescent Imaging for Intravenous Injection

Mice were treated according to approved UK Home Office and institutional guidelines at Imperial College London. Mice were anesthetised with isoflurane using Boyle's apparatus, and injected intraperitoneally with 300µL 15mg/mL D-luciferin salt. Mice were left for 5 minutes exactly and then imaged for bioluminescence by the IVIS Imaging 50 Series (Xenogen) system Caliper Life Sciences Ltd, Runcorn, UK). After acquiring a gray-scale photograph a bioluminescence image was obtained with a 12-cm field of view, a binning (resolution) factor of 1 or 8, and a 1/*f* stop and open filter. Signal intensities were calculated with Livingimage 2.50 software (Xenogen) and expressed as photons per second per centimetre squared per seradian (photons/second/cm_2/sr).

2.2.5.6.1 *Ex Vivo* Luciferase Assay

Quantification of luciferase protein for *ex vivo* tissue homogenates was carried out using the Luciferase Assay System (promega) as per manufacturers instructions. Tissue was freeze-thawed in 1X Lysis buffer, homogenized, and cell debris removed by centrifugation. 20µL of homogenate is added to a 96-well, clear/flat bottom plate. 100µL of luciferase mix is then added to each well and the resulting light emission measured using a luminometer for 30 seconds and reported in relative luciferase units (RLU).

Homogenate is then diluted 1:250 and a Bradford protein assay (BioRad) carried out as per the manufacturers instructions in a clear 96 well microplate to determine the protein concentration of each homogenate in µg/mL. Luciferase quantification in tissue can then reported as relative luciferase units per µg protein (RLU/ µg protein).

2.2. 5. 7 Cellular Staining

Analysis of animal samples of peripheral blood, bone marrow and speenocytes was conducted by staining the sample with appropriate antibodies.

Collected peripheral blood sample (50µl) or $5x10^5$ bone marrow cells (upon animal sacrifice) was washed with 1ml of Immunofluorescence buffer and centrifuged for 5min at 1500rpm. To reduce background binding, and increase mouse specificity of the antibody, cells were incubated with 2.5µl of Fc Block away from light for 10min at 4^oC followed by an IF buffer washing step and resuspension (100µl).

Samples were then divided and incubated with 2.5µl of 200ng/µl control isotype antibody (PE, APC or FITC as appropriate) or 200ng/µl of specific antibody (Ter119-APC, CD3-PE (T cells), CD19-PerCP (B cells), CD11b-APC (monocytes) for 30min, again, away from light at 4^oC.Individual samples were then washed with 1ml IF buffer and fixed in 300µl PBS containing 4% of formaldehyde (v/v) for FACS analysis.

Chapter Three

COMPARATIVE STUDY TO ACCESS THE ABILITY OF BOTH UBIQUITOUS AND TISSUE SPECIFIC ELEMENTS TO EXPRESS WITHIN AN INTEGRATION DEFECTIVE LENTIVIRAL VECTOR CONTEXT

3.1. Aims of Chapter Three

- Design new ID- LVs containing tissue specific regulatory sequences
- Compare expression of newly designed ID-LVs with ubiquitous expressing constructs
- Compare tissue specificity in an ID context, namely globin and muscle specific vectors

3.2 Introduction

Efficient and sustained transgene expression has been previously been reported in non-mitotic cells with use of non-integrating lentiviral vectors *(*Yanez-Munoz *et al., 2006;* Apolonia *et al.,* 2009*)*. These vectors offer advantages over currently used integration proficient vectors as they avoid side effects caused by genome integration namely insertional mutagenesis and its consequent potential cell transformation.

Reports to date have shown that expression from an integration defective vector can be achieved with ubiquitous viral promoters such as that of spleen forming focus virus (SFFV) and the human cytomegalovirus (CMV). It is, however, still unclear whether genomic promoters of either housekeeping or tissue specifically expressing genes can function from within an integration defective context.

In this study we conducted a comparative analysis to access the ability of both ubiquitous and tissues specific genomic regulatory elements to express in an integration deficient LV context.

As a ubiquitous expressing promoter we employed the generic SFFV promoter (Demaison *et al.,* 2002); and as examples of tissue specific control elements we employed:

a) For potential gene therapy for the haemoglobinopathies, a vector combining *HBB* and its control elements;
b) and, for potential muscle gene therapy vectors with transgenes under control of either the human desmin (*DES*) or muscle creatine kinase (*CK-M*) promoter.

3.2.1 The human β-globin gene (*HBB*) promoter and its locus control element (LCR)

HBB is part of a multigene family on the short arm of chromosome 11 with expression restricted to the erythroid lineage. Expression is basal in erythroid progenitor cells but upregulated to very high levels during terminal differentiation, the rate of which is dependent on the upstream hypersensitive (HS) sites of the globin locus control region (LCR; Levings & Bungert, 2002).

The *HBB* core promoter has been narrowed down to approximately 200bp, and consists of several elements, one of which is located 160bp upstream of the transcription start site and which has been shown to increase the efficiency of the promoter after murine erythroleukaemia (MEL) cells have been induced to differentiate (Antoniou *et al.*, 1988). This region contains the (-30) TATA box, (-70) CAAT box and (-90) CAC box motifs necessary for efficient transcription (deBoer *et al.*, 1988).

The *HHB* promoter expression is largely controlled by its potent LCR. The human LCR is located between -5kb and -16kb upstream of the first gene within the *HBB*-like cluster namely β-globin (*HBE*) and comprises 5 regions of extreme DNaseI hypersensitivity (designated as HS1-5). In this study, we used solely the two most transcriptionally potent LCR elements HS2 and HS3 to augment expression from a linked promoter (Miccio et al., 2008).

3.2.2 The human desmin promoter

The desmin gene (*DES*) promoter region used in this study is 1.7kb in length and has been shown to function in all muscle cell types, especially skeletal muscle (Tam et al., 2006; Talbot et al., 2010). DES is an intermediate filament protein, synthesis of which is restricted to cardiac, skeletal and smooth muscle (Li & Paulin, 1993). It is the

first muscle-specific structural protein detected in the heart and somites and its developmental regulation is orchestrated by its promoter in combination with its locus control region (LCR; Raguz et al., 1998; Tam *et al.*, 2006). Desmin expression increases several-fold upon cell differentiation into myotubes (Li & Paulin, 1991).

The *DES* promoter requires a 280bp enhancer for high-level expression located between -973bp and -693bp. This enhancer has been shown to be necessary to confer correct developmental and tissue-specific expression within skeletal muscle. This area consists of three regions, two positive and one inhibitory, which, from analysing expression driven from combinations of these regions in myoblasts and myotubes, have been shown to cooperate to achieve high expression (Li & Paulin, 1993**).**

3.2.3 The muscle creatine kinase (CKM) promoter/enhancer

Muscle creatine kinase (CKM) is a key metabolic enzyme of skeletal muscle and has been shown to be upregulated 17-fold upon muscle differentiation.

The mouse *Ckm* promoter has been extensively characterised. The native, as well as various optimised forms of the promoter/enhancer region have been used to drive skeletal muscle specific expression of both reporter and therapeutic mini and micro-dystrophin genes in the *mdx* mouse model of Duchene muscular dystrophy delivered by both AAV (Salva *et al.*, 2007; Gregorevic *et al.*, 2004) and lentiviral vectors (Li *et al.*, 2005). Mouse *Ckm* is therefore one of the current promoters of choice in the field of muscle gene therapy. In this study we used a 1kb region from -1061bp to +28bp from the human *CKM* transcription start site, which contains the proximal promoter and a 292bp enhancer element between -933 to -641bp that has been shown to confer developmental and tissue specific expression in C_2C_{12} cells (Trask *et al.*, 1988)

3.3 Findings prior to this study

Prior to this project the host lab and collaborators developed a novel LV (GLOBE) where *HBB* is under the control of a minimised promoter and LCR elements HS2 and HS3 (Miccio *et al*., 2008; **Figure 3.1**). This GLOBE LV has been shown to be able to express therapeutic levels of *HBB* in a mouse model of β-thalassaemia (Miccio *et al*., 2008) and correct the α/β-globin chain imbalance in cells derived from patients with β-thalassaemia (Roselli *et al*., 2010).

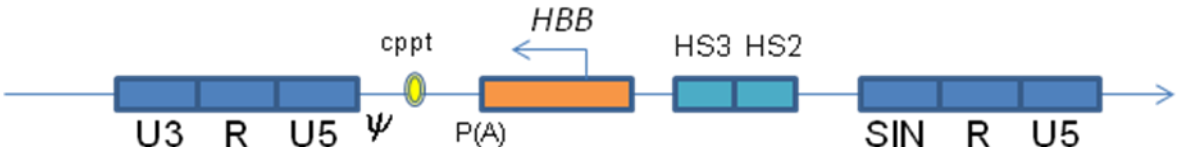


Figure 3.1 - Diagram of the GLOBE LV construct. The construct comprises a minimised *HBB* promoter-gene combination under the control of LCR elements HS2 and HS3. This construct is 9.5kb in length.

Results obtained demonstrate that LCR sites HS2 and HS3 coupled with the prooter-gene combination are sufficient to completely revert the phenotype of severe β-thalassaemia *intermedia* in mice (Miccio *et al*., 2008). The assessment of therapeutic efficacy of this LV was analysed in the β-thalassaemic th3/+ mouse model of this condition, which exhibits chronic anaemia, anomalies in red cell size and shape, splenomegaly, extramedullary hematopoiesis (EMH) and iron accumulation in liver and spleen. Differentiated progeny of th3/+ HSCs, LV-transduced with GLOBE showed a significant increase (79-100%) of donor red blood cells expressing *HBB*, which remained stable up to 9 months after transplantation together with extreme reduction of iron deposition in the spleen and liver (Miccio *et al.*, 2008). Furthermore, the studies performed with GLOBE have shown that the β-thalassaemia correction was achieved by

in vivo selection of genetically corrected erythroblastic progenitors, differentiating from a relatively limited number of transduced HSCs (Miccio *et al.,* 2008).

Further *in vitro* assays were performed to evaluate GLOBE vector efficacy and safety. Using $CD34^+$ cells from β-thalassaemia patients, it was demonstrated that the *ex vivo* manipulation required for successful transduction does not affect cell viability. In addition, at low GLOBE vector copy number per $CD34^+$ cell was sufficient to correct β-thalassaemia globin chain imbalance and restore erythropoiesis upon differentiation *in vitro.* Finally, as is the case with other lentiviral and retroviral vectors, GLOBE integrates preferably within transcriptionally active genes within the target cell population at the time of transduction, with no bias towards integration in the proximity of oncogenes (Roselli *et al.,* 2010).

3.4 Lentiviral Vectors used in this study

3.4.1 Construction of a GLOBE variant containing an eGFP reporter gene

In vitro functional gene analyses are greatly expedited with the easy identification of a transgene transfected into a cell pool. An easy identification requires the incorporation of a readily assayable reporter gene within the vector. Enhanced green fluorescent protein (eGFP) is a protein that aids in the determination of the function of a gene cassette of interest by attaching its coding region to the promoter region, allowing easy detection of transgene expression: the synthesis of eGFP protein when exposed to fluorescent light emits green light. In gene therapy, the presence of a reporter gene in the transfected therapeutic vector allows easy and efficient detection of transduced cells, as these will emit green light when exposed to a UV source.

The minimised *HBB* gene in GLOBE was replaced with a cassette containing the *HBB* promoter driving expression of an eGFP reporter gene linked at the ATG translational start site within exon 1. The removal of the *HBB* mini-gene from GLOBE was achieved using restriction sites situated at either side of the gene; *Cla*I at the 5' end and *Swa*I at 3', both unique sites in the GLOBE vector plasmid. The digestion of this construct with both enzymes released the *HBB* fragment of approximately 1.5kb with the remaining 8kb plasmid vector backbone, containing the βLCR sites HS2 and 3, with the latter to be used for the construction of the new eGFP-containing LV (**Figure 3.2**).

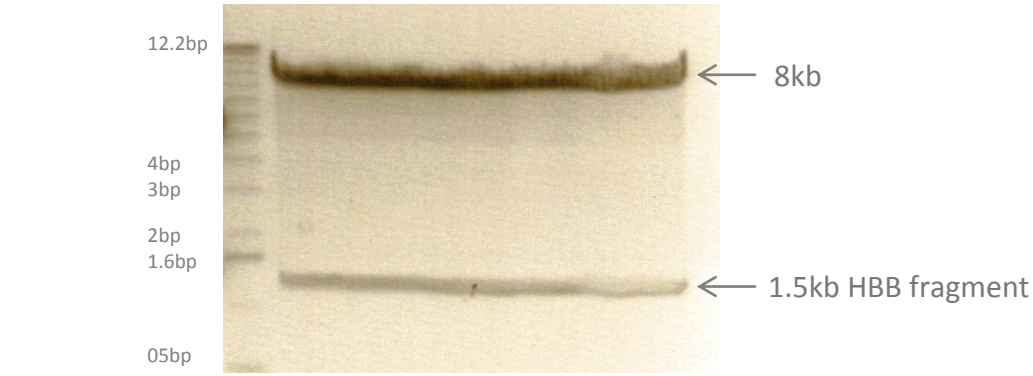


Figure 3.2 - Digestion of GLOBE1 to release HS2 and HS3 loaded in a large well. GLOBE vector digested with *Cla*I and *Swa*I releases a 1.5kb fragment (containing *HBB*) and an 8kb vector backbone containing LCR elements HS2 and HS3. The upper 8kb DNA band was excised from the gel and purified. Invitrogen 1kb DNA size ladder is shown on the left.

A 1.3kb DNA fragment containing the *HBB* promoter linked at the ATG translational start site to an eGFP reporter gene (β-eGFP) was isolated from a plasmid vector available in the host lab (MA891) comprising ~6kb. The β-eGFP fragment was excised as an *Eco*RI (blunted) plus *Cla*I digestion and purified by preparative agarose gel electrophoresis (**Figure 3.3**)

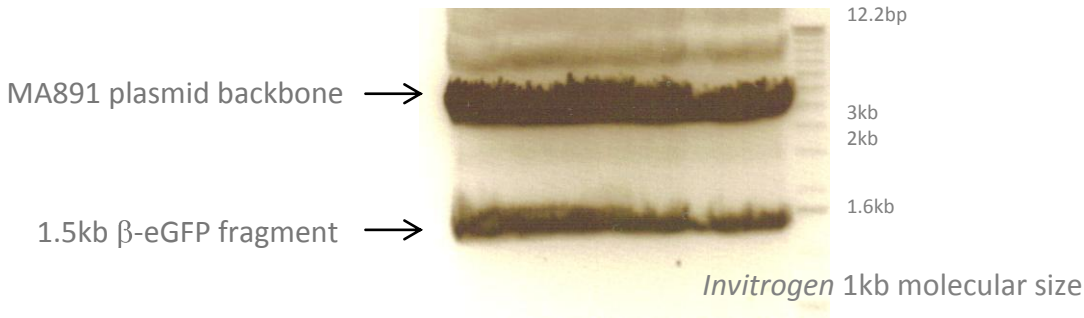


Figure 3.3 - MA891 plasmid was digested with *EcoRI* (blunted) and *ClaI*. Digestion products were resolved by electrophoresis on a preparative agarose gel. The lower band

(1.5kb) comprising the β-eGFP fragment was excised and purified for further cloning. Invitrogen 1kb DNA size ladder is shown on the right.

The β-eGFP and GLOBE vector backbone fragments were ligated employing various insert:vector ratios and reactions used to transform DH5α *E. coli* by the heat shock method and plated to grow on ampicillin agar plates. Individual bacterial colonies were picked and grown overnight in ampicillin-containing LB medium as 5ml mini-cultures. Plasmid DNA was then isolated from a 1ml aliquot of the cultures by the alkaline lysis procedure. The final construct was designated LCR-Globin-eGFP (MA925 under the host lab glycerol system; **Figure 3.4**).

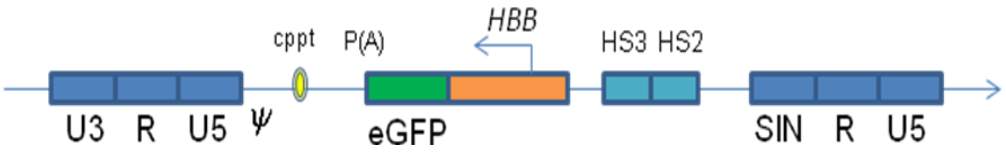


Figure 3.4 - Diagram of the GLOBE-eGFP (LCR-Globin-eGFP) construct. The minimal *HBB* gene was replaced in GLOBE1 (**Figure 3.1**) with a fragment containing the 265bp *HBB* promoter driving expression of a 1.38kb eGFP reporter gene linked at the ATG translational start site.

3.4.2 Design of constructs containing ubiquitous expressing promoters

The spleen forming focus virus (SFFV) promoter driven LV cassette (universally named SEW; **Figure 3.5**) was generated and successfully used in collaborators' laboratories (Demaison *et al*, 2002) and its successful use in ID LV has been documented *(*Yanez-Munoz *et al. 2006;* Apolonia *et al*, 2009*).* We decided to modify this vector and augment its expression in the erythroid lineage with the addition of βLCR sites HS2 and HS3, as present in the GLOBE vector.

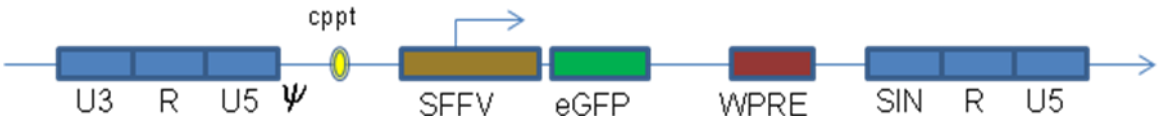


Figure 3.5 - SEW lentiviral vector. The spleen focus forming virus (SFFV) LTR promoter/enhancer drives expression of an eGFP reporter gene. WPRE, the woodchuck hepatitis virus post-transcriptional regulatory element. Arrow denotes direction of transcription.

The Friend spleen focus-forming virus (SFFV) is a highly pathogenic retrovirus that exclusively induces erythroleukemia in susceptible strains of mice within weeks of inoculation (Ruscetti, 1999). A number of studies have been performed using its LTR promoter/enhancer as an internal promoter in LV cassettes as it has been reported that it enhances long-term expression in primary human hematopoietic cells *in vivo* (Demaison *et al*, 2002).

The SFFV promoter is cloned in SEW upstream of the eGFP reporter gene (**Figure 3.5**). The vector also contains the Woodchuck hepatitis virus post-transcriptional regulatory element (WPRE), known to improve vector potency and increase transgene expression levels by facilitating the nuclear export of unspliced RNA to the cell cytoplasm (Zufferey *et al*., 1999).

Direct incorporation of the βLCR HS2 and HS3 elements into SEW by blunt-end ligation proved to be a difficult cloning step. As an alternative strategy we inserted an oligonucleotide polylinker sequence upstream the βLCR HS2/HS3 fragment in a previously designed sub-clone (MA89) to facilitate its isolation and subsequent sticky-end ligation into SEW.

MA89 is a construct with a pBluescript backbone containing βLCR HS2/HS3 elements inserted into in the plasmid's polylinker *Sma*I site, which destroyed upon insertion (**Figure 3.6**).

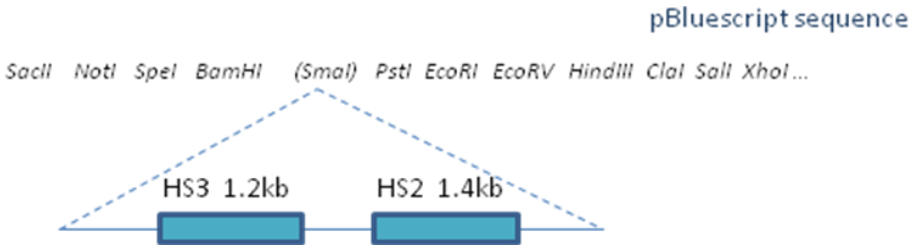


Figure 3.6 – Incorporation of an Oligonucleotide sequence into pBluescript *SmaI* site. The oligonucleotide sequence introduced 5' of HS3 in MA89 at the *Not*I site, contains *Eco*RI and *Xho*I sites, facilitating the removal of the βLCR fragment as either an *Eco*RI or *Xho*I fragment (**Figure 3.7**)

5' GCCCGCTACGTACTCGAGGAATTC 3'

3' CGATGCATGAGCTCCTTAAGCCGG 5'

Not I *Sna BI* *Xho I* *Eco RI* *Not I*

(compatible end) (compatible end)

does not re-form the *Not I* site

Figure 3.7 – The Oligonucleotide sequence

MA89 (6kb) was linearised with *Not* I, whose ends are now compatible with those of the designed oligonucleotide sequence (**Figure 3.7**).

The MA89 vector and the oligonucleotide were ligated and used to transform DH5α *E. coli* by the electroporation. Individual colonies were picked and grown as 5ml mini-cultures overnight. Plasmid DNA from the cultured cell colonies was isolated by the alkaline lysis extraction procedure and subjected to diagnostic restriction enzyme digestion using *Eco*RI and products resolved by agarose electrophoresis (**Figure 3.8**).

NEB 1kb molecular size marker **(A)**

(B)

Figure 3.8 – Digestions of mini preparatios obtained from the MA89/oligonucleotide ligation. *Eco*RI was used to release the βLCR insert if oligonucleotide insertion was successful. **(A)** Samples 1 and 6 show the desired 2.7kb band containing the βLCR HS2 and HS3 sites. NEB 1kb ladder is shown on the far right lane.(B) Diagram showing that the successful insertion of the polylinker oligonucleotide would give rise to a band of 2.7kb (size of βLCR HS2-HS3 sites).

Plasmid from colonies 1 and 6 showed the desired 2.7kb βLCR HS2-HS3 fragment (**Figure 3.8**) and were chosen for further analysis.

Figure 3.9 - Analysis of colonies 1 and 6 obtained from MA89/oligonucleotide ligation. (A) Lines 1, 2 and 5 refer to colony 1 (blue) digested with *Not*I (1), *Xho*I (2) and *Sna*BI (5). Lines 3, 4 and 6 refer to colony 6 (green) digested with *Sna*BI (3), *Xho*I (4) and *Not*I (6). Lines 7, 8 and 9 refer to the original MA89 vector (red) digested with *Xho*I (7), *Sna*BI (8), and *Not*I (9). **(B)** Diagram explaining fragments obtained upon digestion with *NotI*, *SnaBI* and *XhoI*.

Confirmation tests were made by digesting the individual colony DNA with *Not*I, *Xho*I and *Sna*BI, the remaining sites present in the inserted oligonucleotide polylinker. Digestion with these showed positive results for the presence of the oligonucleotide polylinker and βLCR HS2 and HS3 sites in the new vector (**Figure 3.9**) with isolation of the 2.7kb fragment by *SnaBI* (lines 3 & 5) and plasmid division into 2 3kb fragments by *XhoI*. Please note that colony 6 is partially undigested. The oligonucleotide sequence provides new restriction sites namely *Eco*RI, *Sna*BI and *Xho*I 5' of HS3 in MA89. The new construct was designated MA954 in the laboratory glycerol stock system.

3.4.2.1 Incorporation of the βLCR HS2-HS3 elements into the SEW vector

The SEW plasmid was linearised with *Eco*RI and treated with calf intestinal phosphatase (CIP) to avoid vector re-ligation. The vector was then purified and precipitated ready for the insert take up.

The βLCR fragment was removed from the MA945 construct via *Eco*RI digestion thereby releasing the whole HS2-HS3 βLCR fragment (**Figure 3.10**).

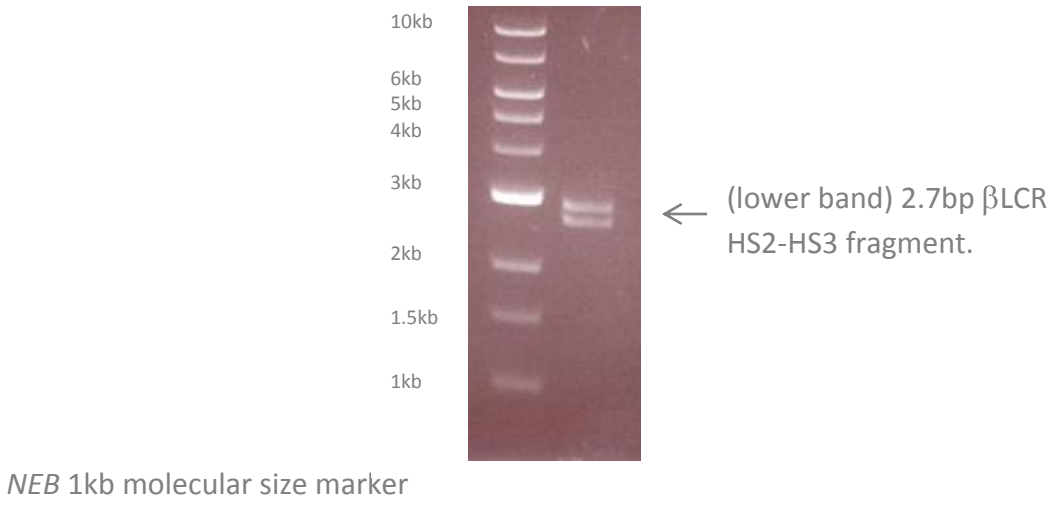


Figure 3.10 - Digestion of MA954 with *EcoRI*. The upper band is the pBluescript vector and lower band the 2.7kb βLCR HS2-HS3 fragment.

The initial *Eco*RI digestion of MA954 (**Figure 3.10**) was followed by a *Sca*I digestion. This cuts within pBluescript and allows clearer separation of the βLCR and plasmid vector backbone after digestion with *Eco*RI. The digestion products were then resolved by agarose gel electrophoresis and the desired band excised and purified.

A ligation reaction of the *Eco*RI linearised SEWs and βLCR HS2-HS3 fragment was performed using a 3:1 insert (βLCR) to vector (SEW) ratio. The ligations were

transformed into electrocompetent LDαGOLD *E coli* cells and spread on agar plates containing ampicillin and left to incubate overnight at 37ºC.

Colonies obtained were used to inoculate 5ml mini-cultures and incubated overnight. The mini-scale cultures were then subjected to alkaline lysis extraction of plasmid DNA. Plasmid DNA obtained from each culture of cell clones underwent appropriate diagnostic restriction enzyme digests and subsequent agarose electrophoresis analysis as follows.

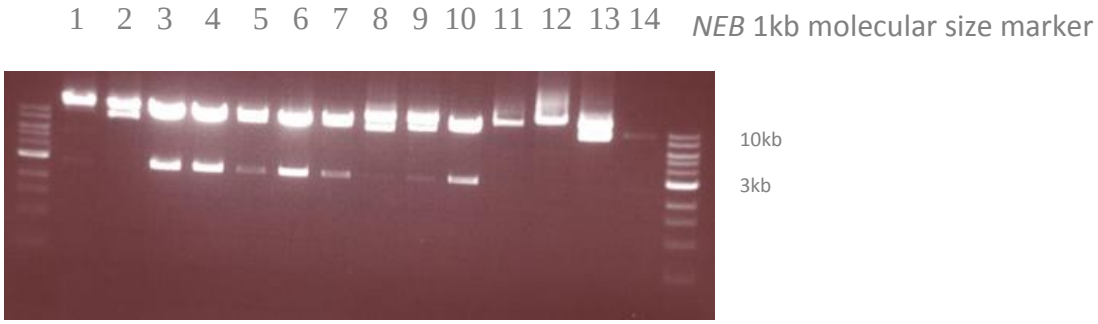


Figure 3.11 - Digestion of plasmid mini-preps obtained from SEW/βLCR HS2-HS3 fragment ligation with *EcoRI*. Mini-preps 3, 4, 5, 6, 7 and 10 were potentially correct giving the expected 2.7kb βLCR HS2-HS3 fragment and were selected for further analysis.

The plasmid mini-preparations were digested with *Eco*RI, which if correct would release the 2.7kb βLCR fragment. On agarose gel analysis colonies 3, 4, 5, 6, 7 and 10 were potentially correct (**Figure 3.11**) and were selected for further analysis.

Selected individual plasmids preparations that gave a positive result from the *Eco*RI digestion (**Figure 3.11**) were digested respectively with *Age*I (to double-confirm the presence of the insert) and *Bam* HI, which determines the insert orientation. AgeI disgests both the plasmid and the insert (A lanes, **Figure 3.12**) *Bam*HI recognises a sequence in the original vector immediately 3'after the βLCR HS2-HS3 fragment insertion and in the insert at its 5' end, in the polylinker just before HS3 (HS2/HS3

diagram, **Figure 3.6**). If the insert is placed in the original vector in the correct orientation; that is, resulting in the order HS3-HS2-SFFV promoter, the *Bam* HI digestion releases the βLCR fragment (lanes 3B, 4B and 6B of **Figure 3.12**). On the other hand, if the insert has ligated in the reverse orientation, the two recognition sequences will be placed very close to each other and only a 12 kb fragment will be visible upon the gel electrophoresis analysis (lanes 5B, 7B and 8B of **Figure 3.12**)

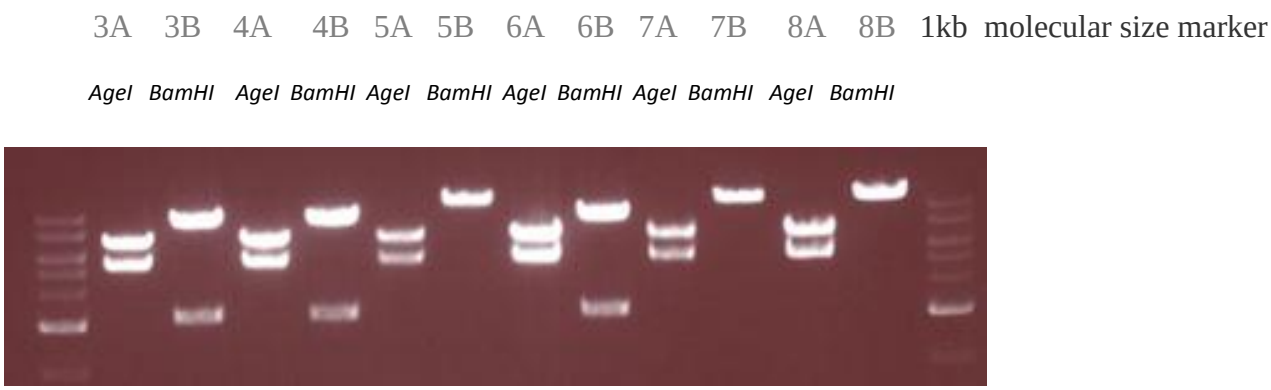


Figure 3.12 - Analysis of plasmid mini-prep numbers 3, 4, 5, 6 , 7 and 10 from SEW/βLCR ligation (Figure 3.11). Plasmids were digested with either *Age*I **(A)** or *Bam*HI **(B).**

Plasmids from colonies 3, 4 and 6 show the correct diagnostic digests (**Figure 3.12**). Colony 4 was selected for further use and this new construct was designated as **βLCR-SEW** (MA945 in the host lab glycerol stock)**.**

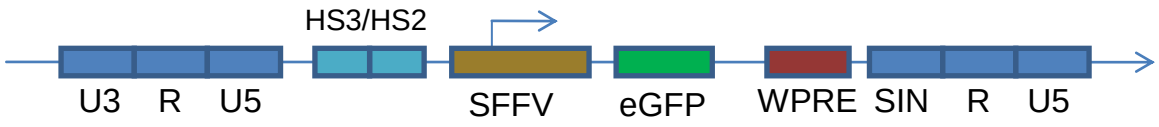


Figure 3.13 - Diagram of the βLCR-SEW vector. The βLCR together with the SFFV promoter driving expression of an eGFP reporter gene.

3.4.3 Addition of another *HB* construct into this study

As described previously (section **1.2.5**), studies have shown that intron 2 of *HBB* is essential for promoting efficient mRNA 3'end formation (cleavage and polyadenylation) (Antoniou *et al.*, 1998*;* Millevoi *et al.*, 2002). With this in mind we decided to incorporate in this study a vector where the eGFP reporter gene is inserted between exons 1 and 2 of *HBB*. Although this HBB-eGFP cassette lacks parts of exons 1 and 2 as well as intron 1, it crucially retains intron 2. Therefore, this transcription unit consists of the following components; βLCR HS2-HS3 sites, 265bp *HBB* promoter, 3' half of *HBB* (partial exon 2, intron 2 and exon 3) and eGFP cDNA sequence inserted between exons 1 and 2 (**Figure 3.14**). This HBB-eGFP vector (MA951) was generated by *Vincent Kao* in the host laboratory.

Generally, this expression vector design has been demonstrated to express a range of heterologous protein types rapidly and easily, including secreted products (Needham *et al.*, 1992 & 1995).

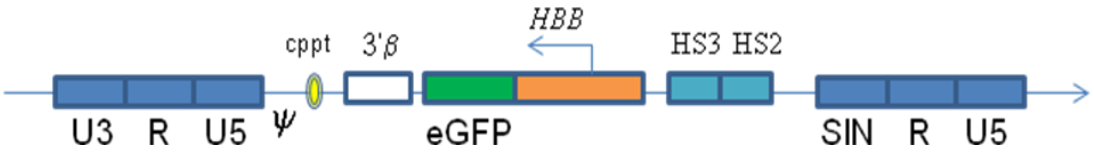


Figure 3.14 - Diagram of the *HBB*-eGFP vector (MA951). Vector courtesy of *Vincent Kao*, a member of the host laboratory)

3.4.4 *CK-M* and *DES* constructs

Expression from ubiquitous regulatory elements from within an ID-LV context has been extensively documented (Apolonia *et al*., 2009; Yanez-Munoz *et al.*, 2006). However, there is a lack of knowledge with regards to expression from tissue specific elements within this system.

We therefore decided to broaden our study to also test muscle specific vectors as well as erythroid *HBB* cassettes in an ID-LV configuration. LVs containing respectively human *DES* and *CK-M* promoter-enhancer elements driving expression of an eGFP reporter gene that were already available in the host laboratory having been produced by *Gillian Talbot* (**Figure 3.15)**.

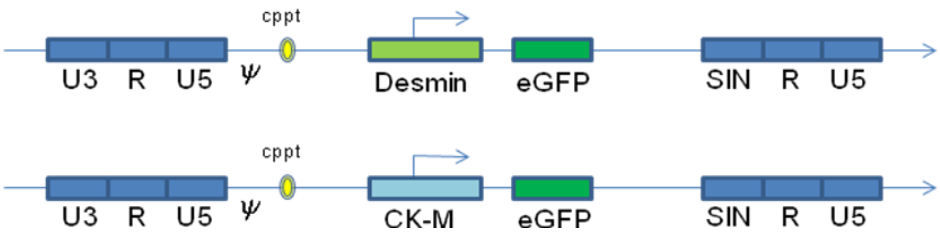


Figure 3.15 - Desmin and CK-M LV vectors. (**A**) SIN LV with an internal Desmin promoter driving eGFP. (**B**) SIN LV with internal CK-M promoter driving eGFP. Both constructs are courtesy of *Gillian Talbot*, a former PhD student in the host laboratory.

3.5 Analysing *HBB* tissue specificity in an integration deficient LV context

This study was performed in K562, a human myelogenous leukaemia cell line (erythroleukemia type) derived from a 53 year old female patient with chronic lymphocytic leukaemia. This is a robust line of suspension cells with recent studies showing that K562 blasts are multipotential, hematopoietic malignant cells that spontaneously differentiate into recognisable progenitors of the erythrocyte, granulocyte and monocytic series. However, these cells do not require differentiation in order to display an erythroid character as they constitutively express embryonic ε-globin and foetal γ-globin, making them a good test system for this experiment.

3.5.1 Lentiviral vector preparation and transduction of cell lines

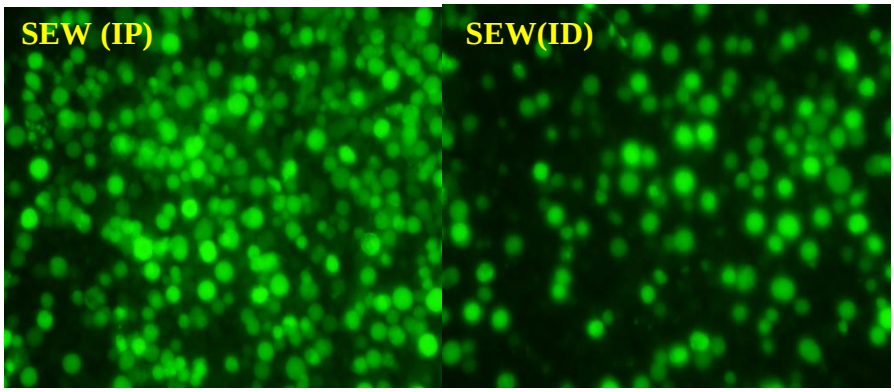
Lentiviral vectors were produced by transient co-transfection of HEK293T cells with 3 plasmids (the lentiviral vector, pMD.G2 [envelop plasmid], and pCMVΔ8.91 [packaging plasmid] employing polyethylenimine (see section **2.2.3**). Viral vector titre of eGFP-containing preparations was determined by transducing K562 cells with serial dilutions of the viral stock and monitoring expression after 3 days by flow cytometry analysis.

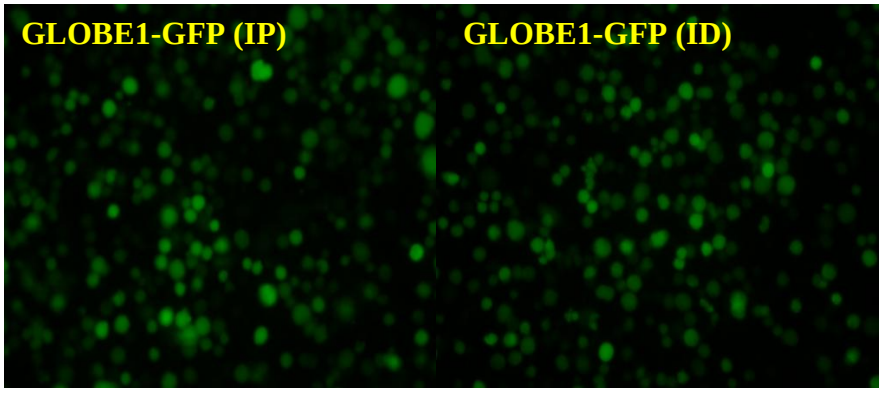
3.5.2 Transduction of K562 cells with integration proficient and integration deficient LV constructs

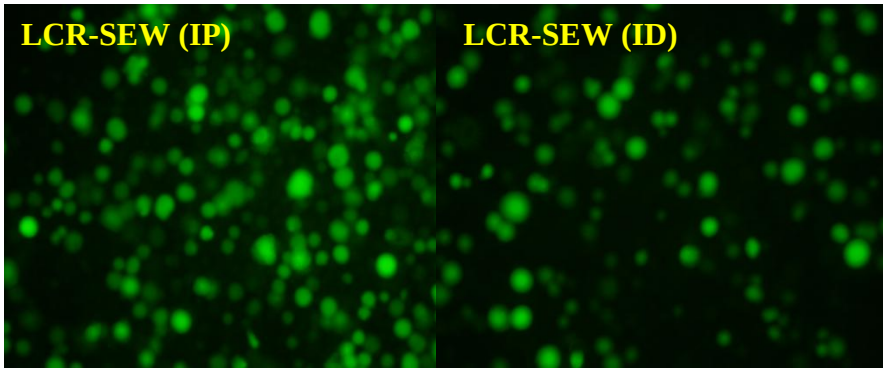
Lentiviral vectors containing eGFP under control of the different *HBB*-based combinations and SFFV promoters were used to transduce K562 cells at a multiplicity of infection (MOI) of 3 in an attempt to achieve transduction efficiency of 1 viral copy per

cell and therefore minimise the frequency of cells harbouring multiple integrations. Pictures of trasduced pools taken under fluorescent light emission microscope can be found in **Figure 3.14**.

A total of 3 pools of transduced cells per construct were generated with identical amounts of virus to control for experimental reproducibility. Cells were collected and analysed for eGFP expression by flow cytometry at 3, 6 9 and 15 days post-transduction. In addition, cells were harvested at the same time points and DNA isolated for determination of average vector copy number (VCN) per cell.


SEW (IP)
SEW(ID)


GLOBE1-GFP (IP)
GLOBE1-GFP (ID)


LCR-SEW (IP)
LCR-SEW (ID)

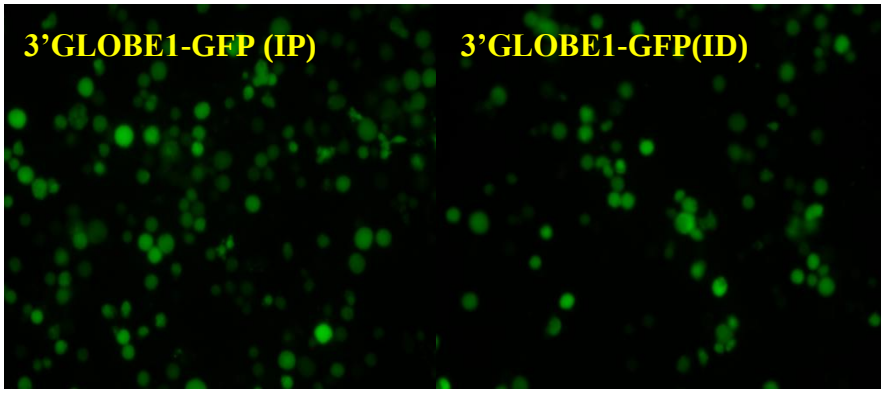

3’GLOBE1-GFP (IP)
3’GLOBE1-GFP(ID)

Figure 3.14 - Images of GFP expression in K562 cells transduced with different LVs. Capture of the eGFP expression of K562 cell pools transduced in both integration proficient and deficient configuration. The pools were transduced with an equal initial viral amount (MOI of 3). Exposure: t = 7s, f/4. Magnification: SEW & LCR-SEW (ID & IP) x20, GlOBE1-GFP & 3'GLOBE1-GFP (ID & IP) x!0.

3.5.3 Comparison of percentage of eGFP positive cells

Figure 3.15, shows the percentage of eGFP positive cells in the pools of transduced K562 cells with the exclusive HBB-based vectors (GLOBE-GFP and HBB-eGFP) and the SFFV promoter driven vectors (LCR-SEW and SEW) in both integration proficient (IP) and deficient (ID) vector configurations.

The vectors in IP configuration showed a decrease in the number of eGFP positive cells by 3-fold over a period of 21 days, with all but the LCR-SEW reaching a plateau stage in the last days of the experiment. All 3 sample wells showed similar values with small variability (**Figure 3.15,** upper panel).

The percentage of eGFP positive cells transduced with the ID vectors also decreased over time but on this occasion reached values close to zero by the end of the 21-day period of the experiment (**Figure 3.15,** lower panel). A decrease in eGFP positive cells was evident with all constructs starting from between 3 and 6 days post-transduction. However, in the case of both SFFV driven vectors a higher level of positive cells was maintained over the last two time points, even though the error bars in the ID graph (**Figure 3.15,** lower panel) are larger when compared with those in the IP analysis (**Figure 3.15,** upper panel), implying more variability between well samples.

The initial drop in eGFP positive transduced cells in ID configuration was expected since these vectors remain in the nucleus as episomal DNA circles and are eventually lost by cellular division (see section **1.3.37**).

3.5.4 Mean Fluorescence Intensity (MFI)

Figure 3.16 shows the different values obtained in Mean Fluorescence Intensity (MFI) for the constructs in both IP and ID configurations. The MFI measures the spread in fluorescence intensity of a given population of cells.

In the IP-LV experiment (**Figure 3.16,** upper panel) the population of positive cells retained expression at the same intensity throughout the experiment until the last analytical point. This consistent expression for all constructs can be interpreted as indicating no changes in active eGFP reporter gene expression from the transduced population of cells.

In an ID-LV context (**Figure 3.16**, lower panel) the intensity of expression decreases by an average of 1-fold of the initial MFI.

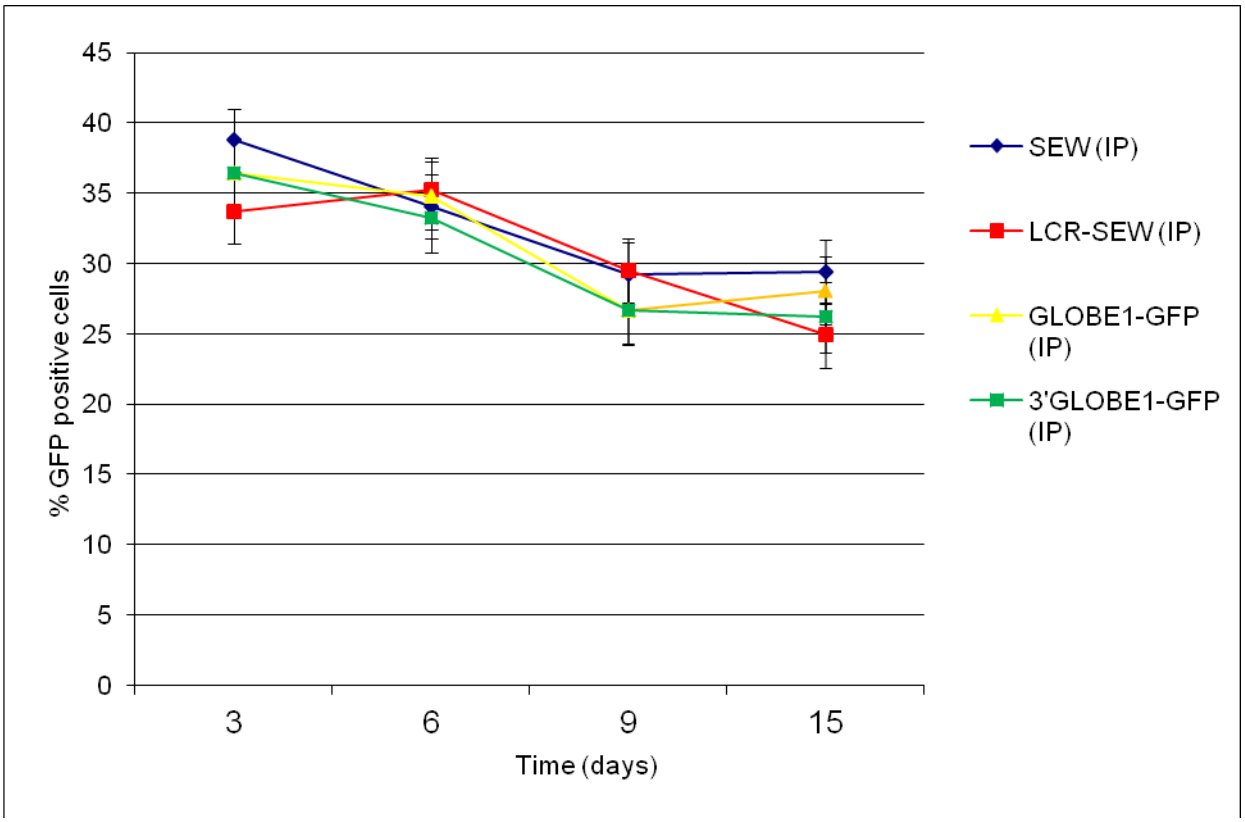
Interestingly, the vectors built as ID express initially with a much lower intensity than those in an IP configuration; at the first time point (day-3) IP vectors express within the range of 20-50 whereas the ID vectors show a range of MFI of 10-20.
This observation raises questions about the feasibility of ID-LVs in the study of globin disorders. For any given therapy a given amount of initial intensity of expression is required – the search for a therapeutic approach for the haemoglobinopathies has shown that at least 30% of the vectors are needed to initial express.

3.5.5 Coefficient of Variegation (CV)

The coefficient of variegation (CV) is a normalised measure of dispersion of a probability of distribution of a given sample population. This accounts for the variation between individual cells in a given cell pool; that is, how much does each gated cell differ from another in level of expression.

The CV measurement only makes sense in positively transduced cells, hence why it's only possible to generate a graph with the IP-LV constructs (**Figure 3.17**). The results show that both constructs driven by a tissue specific promoter (*HBB*) interact with the *HBB* regulatory elements, causing sample variation. However, when a ubiquitous expressing promoter (SFFV) is driving expression (SEW and LCR-SEW) this variation is

reduced. Interestingly, when HS2 and HS3 elements of the *HBB* LCR are added to SFFV the variation between samples diminishes considerably (LCR-SEW), which suggests that the inclusion of an *HBB* element accounts for the consistent expression and reproducibility. Note that the error bars of each sample are relatively small.


45
40
35
30
25
20
15
10
5
0
% GFP positive cells
3
6
9
15
Time (days)
SEW (IP)
LCR-SEW (IP)
GLOBE1-GFP (IP)
3'GLOBE1-GFP (IP)

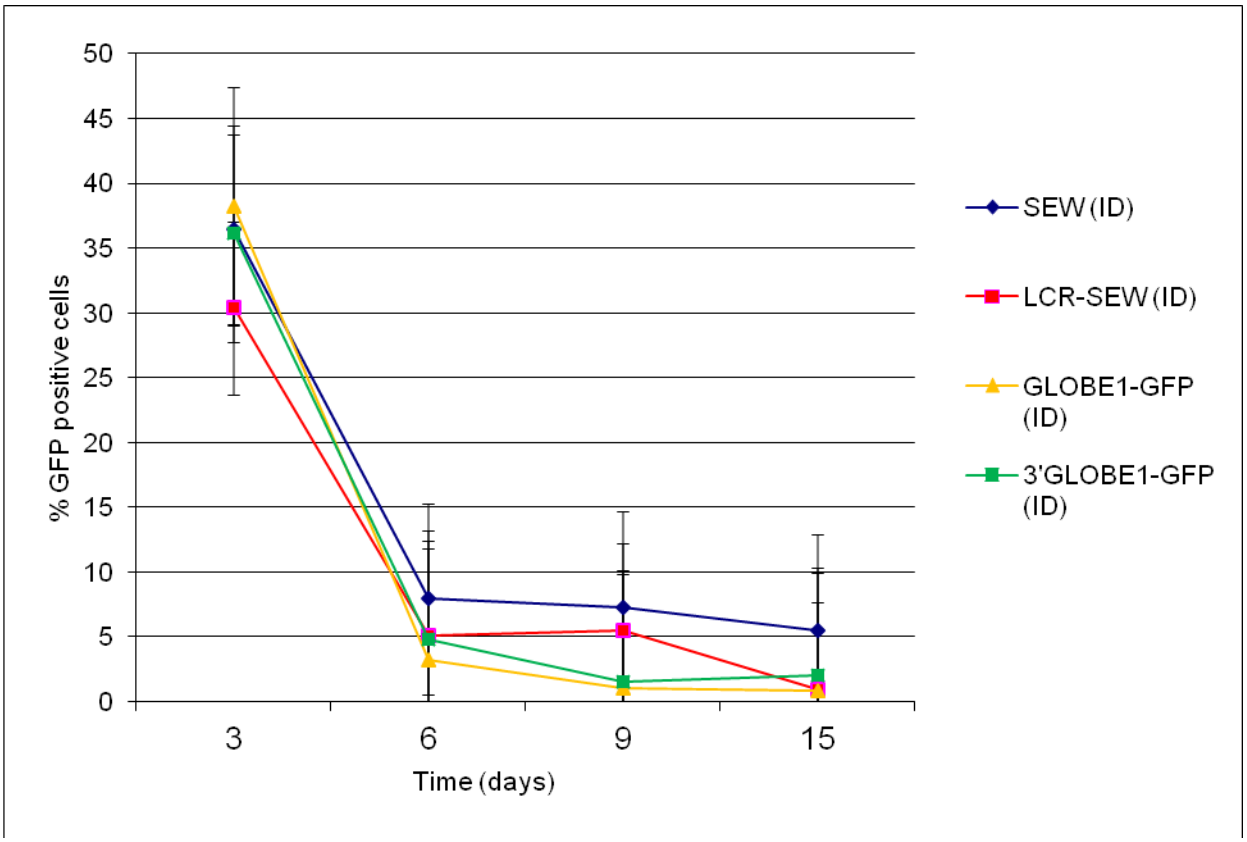
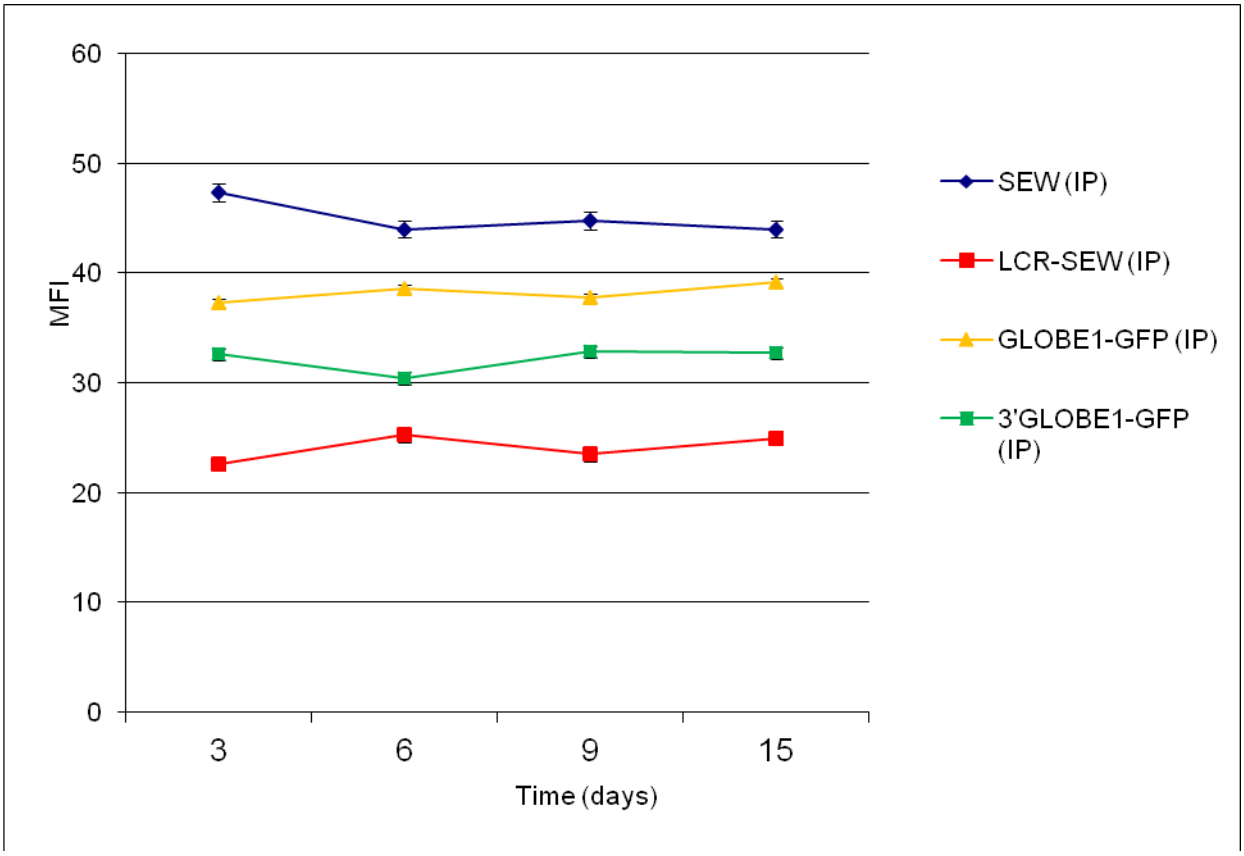

50
45
40
35
30
25
20
15
10
5
0
% GFP positive cells
3
6
9
15
Time (days)
SEW (ID)
LCR-SEW (ID)
GLOBE1-GFP (ID)
3'GLOBE1-GFP (ID)

Figure 3.15 - Percentage of K562 GFP positive cells in cell pools transduced with both IP (upper panel) and ID (lower panel) LV vectors at MOI 3 over a period of 21 days. eGFP expression was determined by flow cytometry of K562 cell pools harbouring integration proficient (IP) and deficient (ID) lentiviral vectors containing and eGFP reporter driven by HBB, 3' globin HBB and SFFV promoters. Points in the graph represent an average obtained for 3 pools (n=3) and the error bars the standard error of the mean.

Statistical analysis was performed using one way ANOVA to detect statistical differences between population means. The decrease of expression in ID vectors (timepoint 3 to 6days, lower panel) is highly significant when compared with the mean population in the same graph at different timepoints (P-value <0.001). Furthermore, the comparison of population means both in IP and ID graphs shows that the decrease of expression in ID vectors is statistical significant (P-value <0.05).


60
50
40
30
20
10
0
MFI
3
6
9
15
Time (days)
SEW (IP)
LCR-SEW (IP)
GLOBE1-GFP (IP)
3'GLOBE1-GFP (IP)

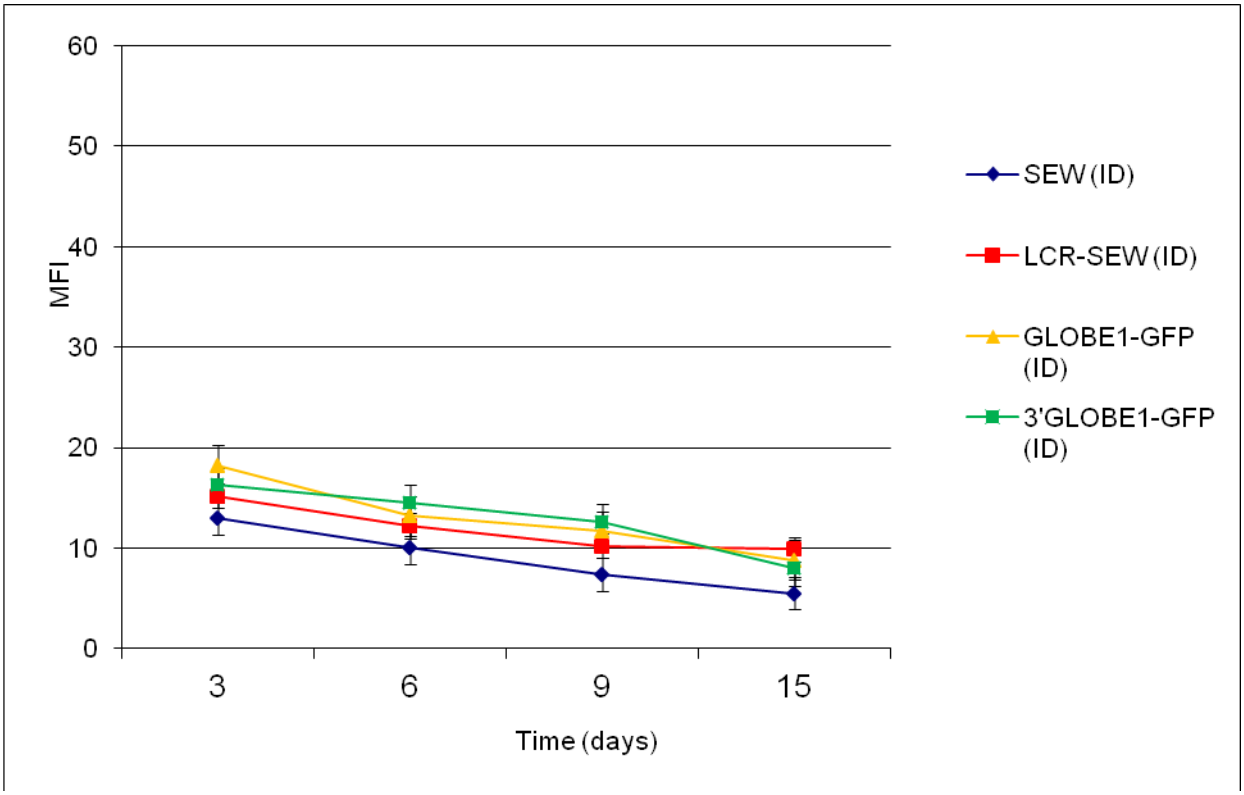

60
50
40
30
20
10
0
MFI
3
6
9
15
Time (days)
SEW (ID)
LCR-SEW (ID)
GLOBE1-GFP (ID)
3'GLOBE1-GFP (ID)

Figure 3.16 - Mean fluorescence intensity (MFI) of K562 cells transduced with both IP (upper panel) and ID (lower panel) LV vectors at MOI 3 over a period of 21 days. MFI shows lower values in ID configuration when compared to IP vectors. However, readily detectable expression from erythroid-specific ID-LV is observed. Points in the graph represent the average obtained for 3 pools (n=3) and the error bars the standard error of the mean.

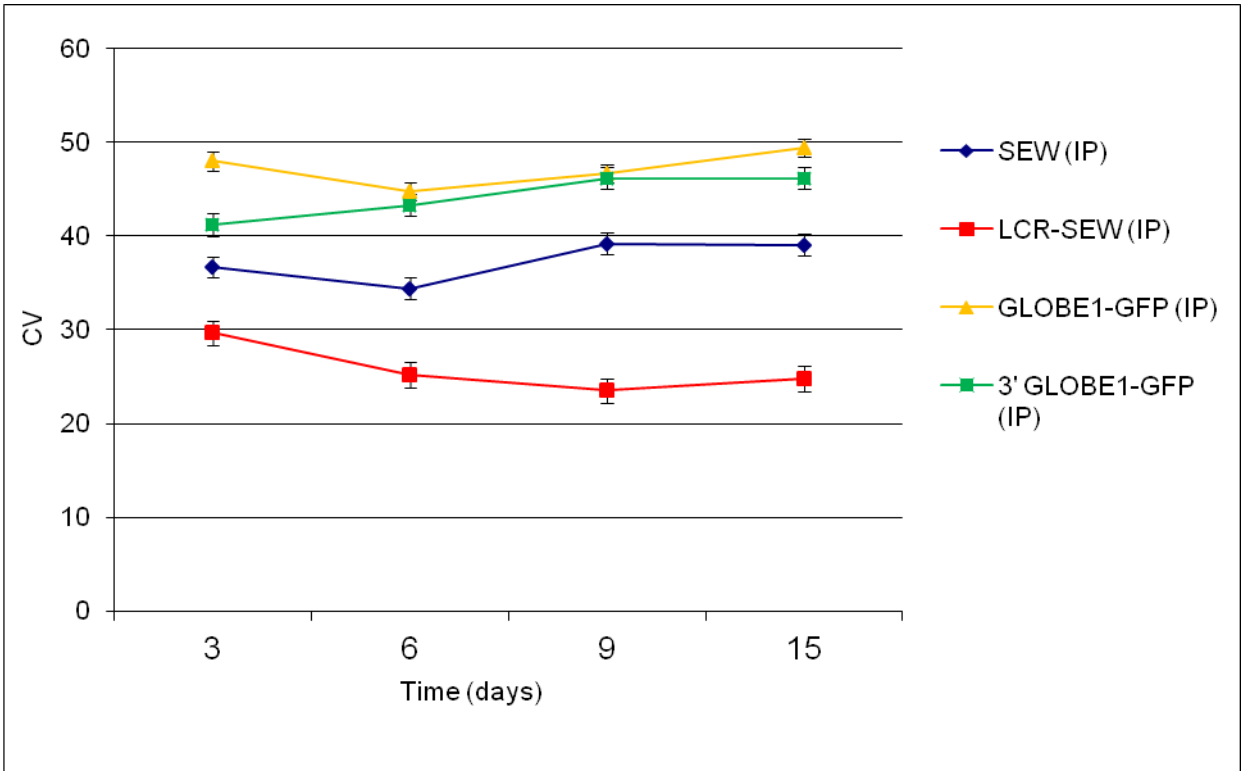


Figure 3.17 - Coefficient of variegation (CV) of eGFP reporter gene expression in K562 cells transduced with IP vectors at MOI 3 over a period of 21 days. CV of expression decreases with the addition of the erythroid-specific βLCR elements to the ubiquitous SFFV promoter. Points in the graph represent an average number obtained for 3 pools over a period of 15 days and the error bars the standard error of the mean.

3.5.6 Vector copy number variation

Determination of average LV copy number per cell of a given sample was only determined once the experiment reach its final stages.

According to the results obtained (**Figure 3.18**), the K562 cells transduced with the IP-LVs containing the *HBB* elements are maintained and accumulate in the cell pool. In contrast the cells transduced with the SFFV vector show a decrease in average copy number per cell over the 15-day period of the analysis.

In an ID-LV context (**Fig. 3.18**, lower panel), as expected the average vector copy number per cell markedly decreases over the 15-day period of the experiment due to dilution resulting from rounds of cell division.

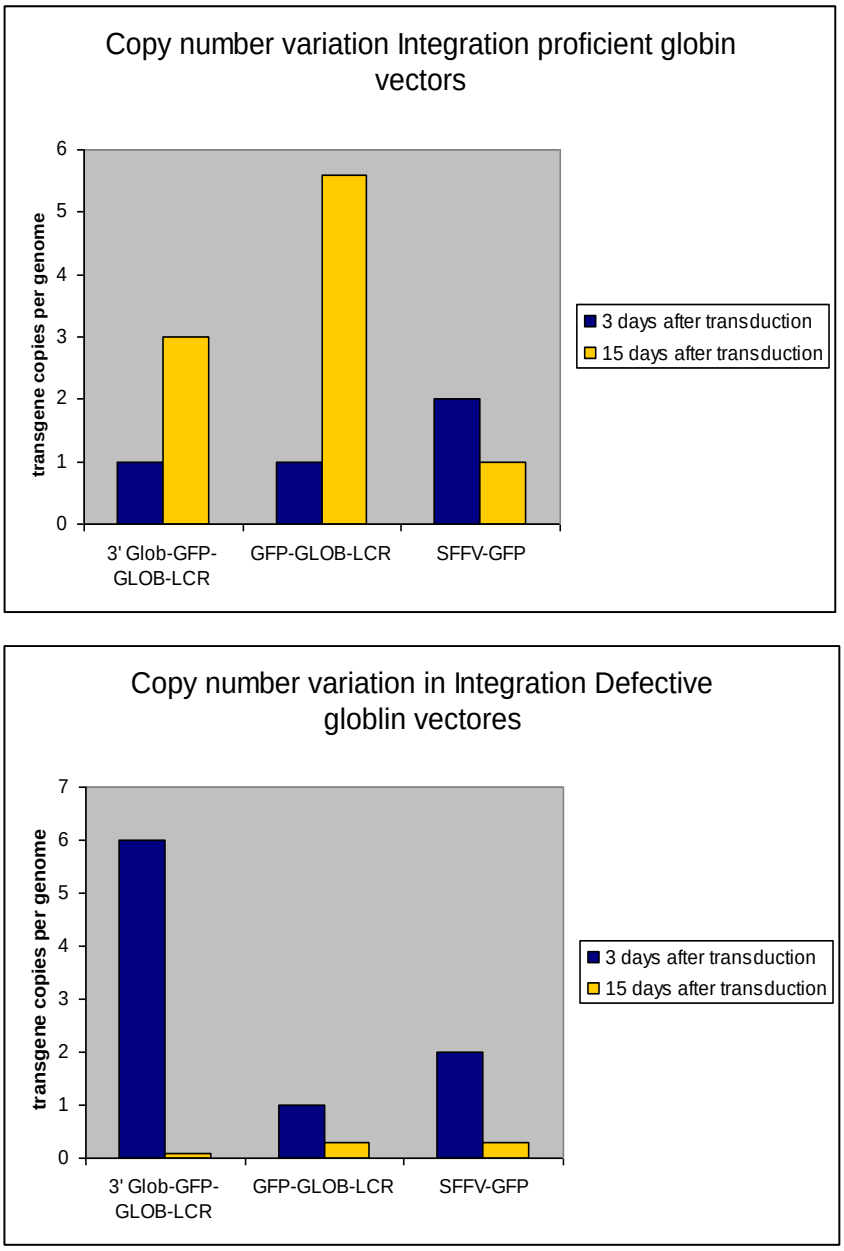


Figure 3.18 - Copy number variation in K562 pools transduced with IP (upper panel) and ID (lower panel) vectors (n=3). DNA extracted from transduced wells was used to determine average LV copy per cell. The copy number increases in an IP configuration in correlation with % of positive cells with an exception for the SFFV driven vector. Vector expression regardless of vector copy number can be explained by the presence of the WPRE element, known to increase transgene expression by facilitating the nuclear export of unspliced RNA to the cell cytoplasm (Zufferey *et al.*, 1999).

3.6 Analysing muscle specificity in an ID-LV context

As referred to previously the evaluation of tissue specific expression from an ID-LV context was extended to include vectors containing muscle-specific regulatory elements. The use of muscle specific vectors requires a different type of cell to be used. In this study we chose to use the murine myoblast cell line C_2C_{12} (Section **2.1.9**, Chapter 2). These are a commonly used model system to study *DES* expression and other muscle specific systems, as the cells can easily be induced to fuse to form terminally differentiated myotubes in order to examine myogenic expression systems.

3.6.1 Lentiviral vector preparation and transduction of C_2C_{12} line

Again, LV vectors were produced by transient co-transfection of HEK293T cells with 3 plasmids (the lentiviral vector, pMD.G2 [envelop plasmid], and pCMVΔ8.91 [packaging plasmid]), employing PEI methodology. Viral vector titre was determined by transducing C_2C_{12} cells with serial dilutions of the viral stock and monitoring expression after 3 days by flow cytometry.

3.6.2 Transduction of C_2C_{12} cells with integration proficient and integration deficient LV constructs

LVs were used to transduce C_2C_{12} cells at MOI 3 in an attempt to achieve transduction of 1 vector copy per cell. A total of 3 pools of cells were transduced with identical amounts of virus per construct to monitor experiment reproducibility. Transduced cells were collected and pools were analysed by flow cytometry at 3, 6 9 and 15 days post-transduction.

DNA isolated from the cell pools was used to determine average vector copy number per cell.

3.6.1 Comparison of percentage of eGFP positive cells

Wells transduced with both muscle-promoter driven vectors (*DES, CKM*) were kept in culture for 15 days and were analysed for eGFP expression at 3, 6, 9 and 13 day time points.

The muscle constructs in an IP configuration showed a relative constant eGFP expression along the time period of the experiment (**Figure 3.19**, upper panel). In an ID configuration (**Figure 3.19**, lower panel), the results show a visible decrease in the percentage of GFP positively expressing cells, again indicating ID vector dilution upon successive cell division.

Both the *DES* and *CK-M* vectors showed the same pattern of expression over time in both in IP and ID configurations. As IP-LV both *DES* and *CK-M* gave expression per sample that decreased only slightly at each time point but with percentage of GFP expressing cells having possibly reached a plateau value by day-15 with the *DES* construct. With ID-LV, both *DES* and *CK-M* vectors suffered an abrupt loss between day-3 and day-6 and then a more gradual decline thereafter. It would be interesting to repeat this experiment for a longer period of time to access if expression with both IP- and ID-LVs continued to decline or remained constant.

Studies with ID-LV configurations in skeletal muscle tissue *in vivo* have been reported previously (Yanez-Munoz *et al*., 2006; Apolonia *et al*., 2009). However, these studies were performed with non-muscle specific ubiquitous expression cassettes. In addition, due to the fact that, unlike tissue culture cells, *in vivo* myofibres are post-mitotic loss of expression was not so pronounced.

3.6.2 Mean fluorescence intensity

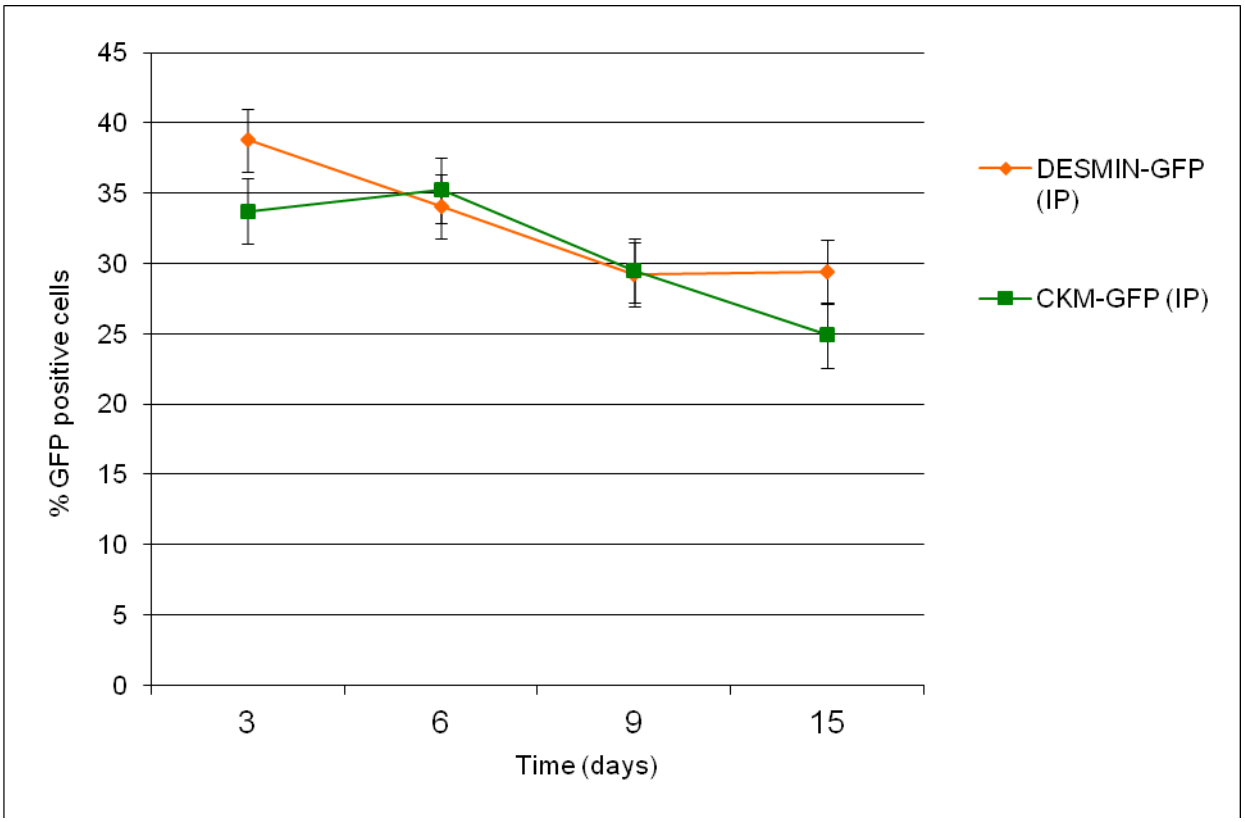
Both *DES* and *CK-M* muscle constructs maintained the initial intensity of eGFP expression throughout the study until the da-15 time point (**Figure 3.20**). However, it is interesting to notice that despite similar profiles in terms of percentage of eGFP positive cells (**Figure 3.19**), actual MFI differed significantly between the 2 muscle vectors in an IP context with *DES* showing a 2-fold higher level of expression than *DES* (**Figure 3.20**, upper panel).

This higher MFI conferred by the *DES* construct over *CK-M* is not due to differences in vector copy number as these values between the two sets of transduced cells were very similar at both 3- and 15-days post transduction

In an ID-LV context the muscle constructs have comparable MFI values throughout the experimental period (**Figure 3.20**, lower panel) with comparable average vector copy number per cell (**Figure 3.22**, lower panel). Again, as noted before with other vectors used in this study, the MFI of a given muscle vector in an ID configuration was always found to be lower when compared with the same vector built in an IP context.

3.6.3 Coefficient of variegation

Figure 3.21 shows that CV values were comparable and consistent over the 21-day duration of the experiment between the *DES* and *CK-M* LVs.


45
40
35
30
25
20
15
10
5
0
% GFP positive cells
3
6
9
15
Time (days)
DESMIN-GFP (IP)
CKM-GFP (IP)

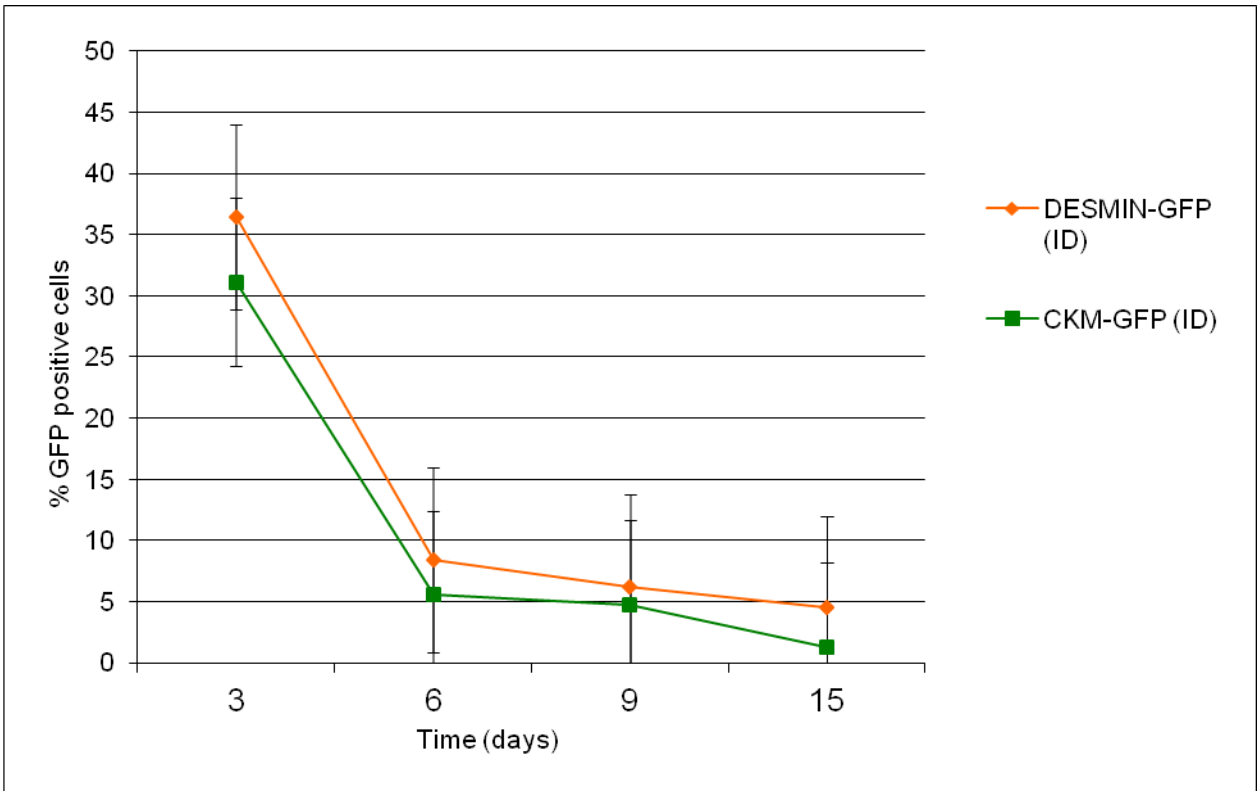
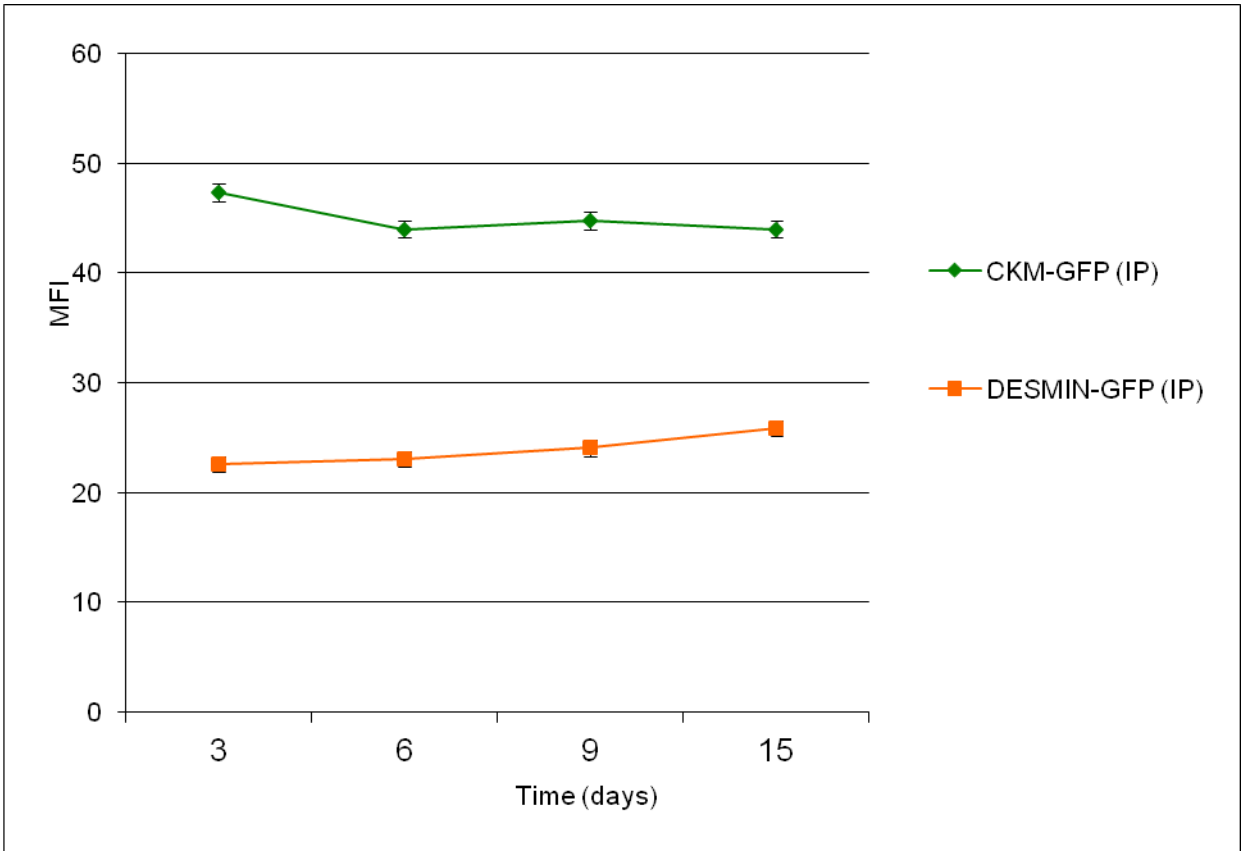

50
45
40
35
30
25
20
15
10
5
0
% GFP positive cells
3
6
9
15
Time (days)
DESMIN-GFP (ID)
CKM-GFP (ID)

Figure 3.19 - Percentage of C_2C_{12} eGFP positive cells in pools transduced with LVs in both an IP (upper panel) and ID (lower panel) configuration at MOI 3 over a period of 21 days. GFP expression was determined by flow cytometry analysis of C_2C_{12} cell pools harbouring lentiviral vectors of both *DES* and muscle creatine kinase (*CK-M*) promoter driven constructs. Again, one way ANOVA to detected statistical differences between population means in an ID and IP context (P-value <0.05). Points in the graph represent an average obtained for 3 pools (n=3) and the error bars the standard error of the mean.


60
50
40
30
20
10
0
MFI
3
6
9
15
Time (days)
CKM-GFP (IP)
DESMIN-GFP (IP)

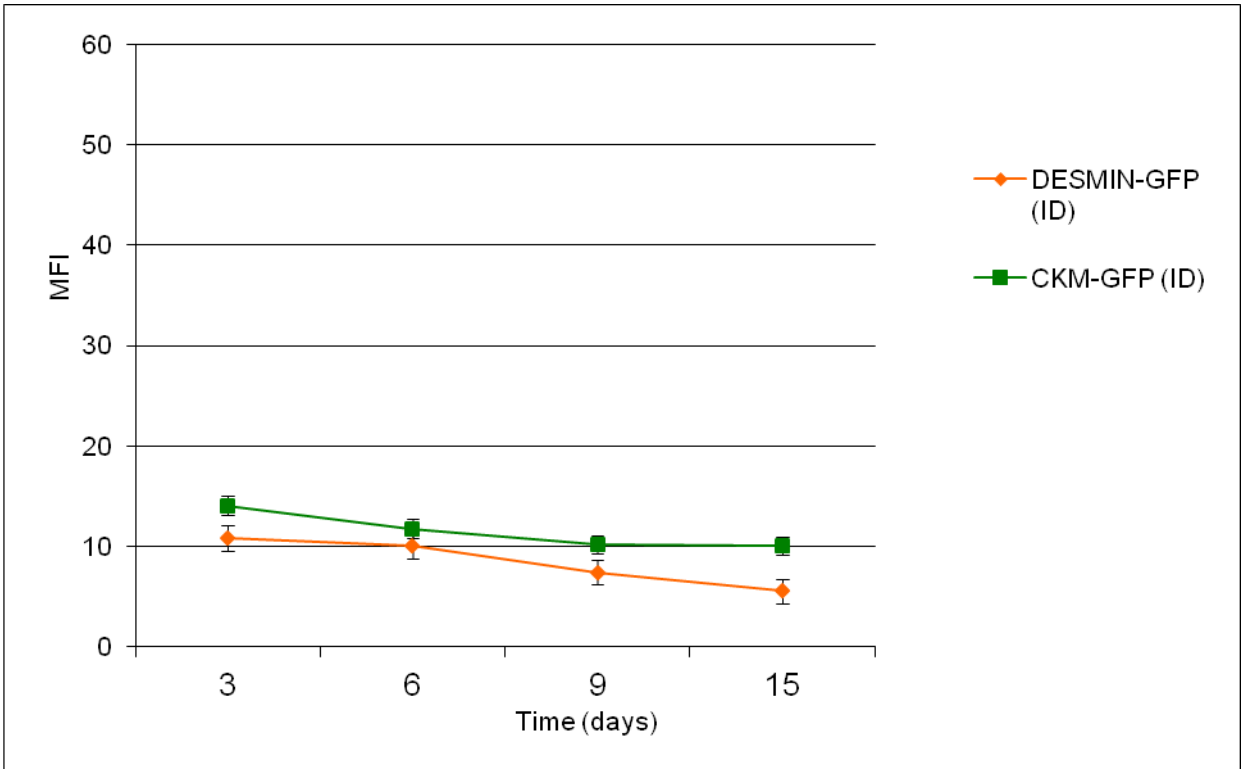

60
50
40
30
20
10
0
MFI
3
6
9
15
Time (days)
DESMIN-GFP (ID)
CKM-GFP (ID)

Figure 3.20 - Mean fluorescence intensity (MFI) of C_2C_{12} cells transduced with Desmin and *CKM* LV in both IP (upper panel) and ID (lower panel) configurations. Cells were transduced at MOI 3 and eGFP expression monitored by flow cytometry at periodic intervals over a period of 21 days. Error bars refer to the variation amongst cell pools.

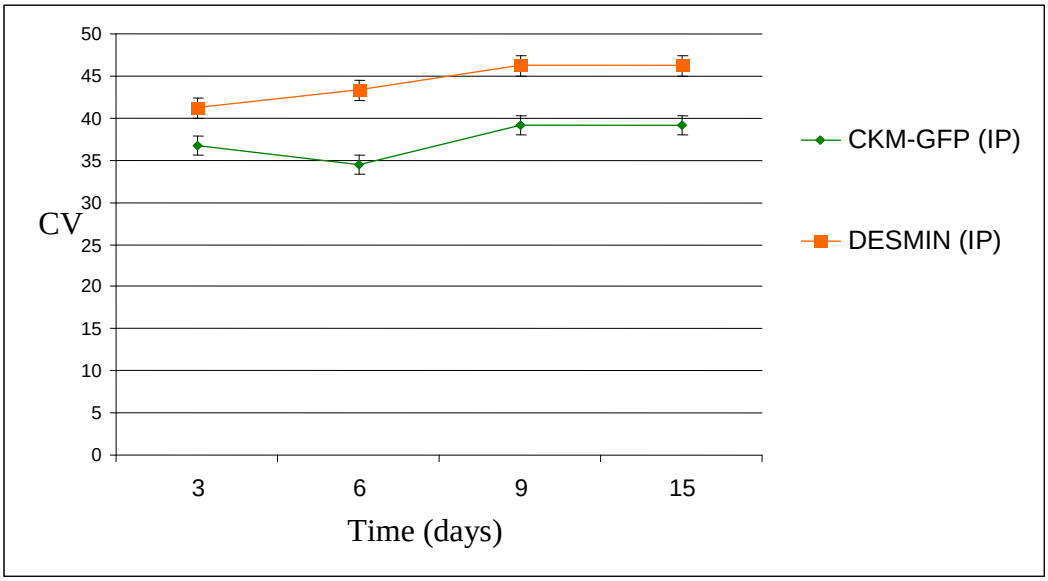


Figure 3.21 - Coefficient of variegation exhibited by C_2C_{12} cells transduced with *DES* and *CK-M* IP-LVs at MOI 3 over a period of 21 days. Error bars refer to the variation amongst cell pools.

3.6.4 Vector copy number

DNA was isolated from cell samples and used to determine average vector copy number per cell at the different time points of the experiment (**Figure 3.22)**. The vector copy number per cell in the IP-LV transductions was similar with both *DES* and *CK-M* constructs and did not alter significantly from initial 3-day time point to the final sample taken at 15-days. Both muscle constructs decreased their amount of vectors per well the upper panel on **Figure 3.22**, shows a vector copy number decrease from 6 to 5 for *DES* and 7 to 5 for *CKM*.

Unsurprisingly, the average vector copy number per cell in samples transduced with ID-LV decreased significantly, as these episomal vectors were expected to be diluted and lost through successive rounds of cell division.

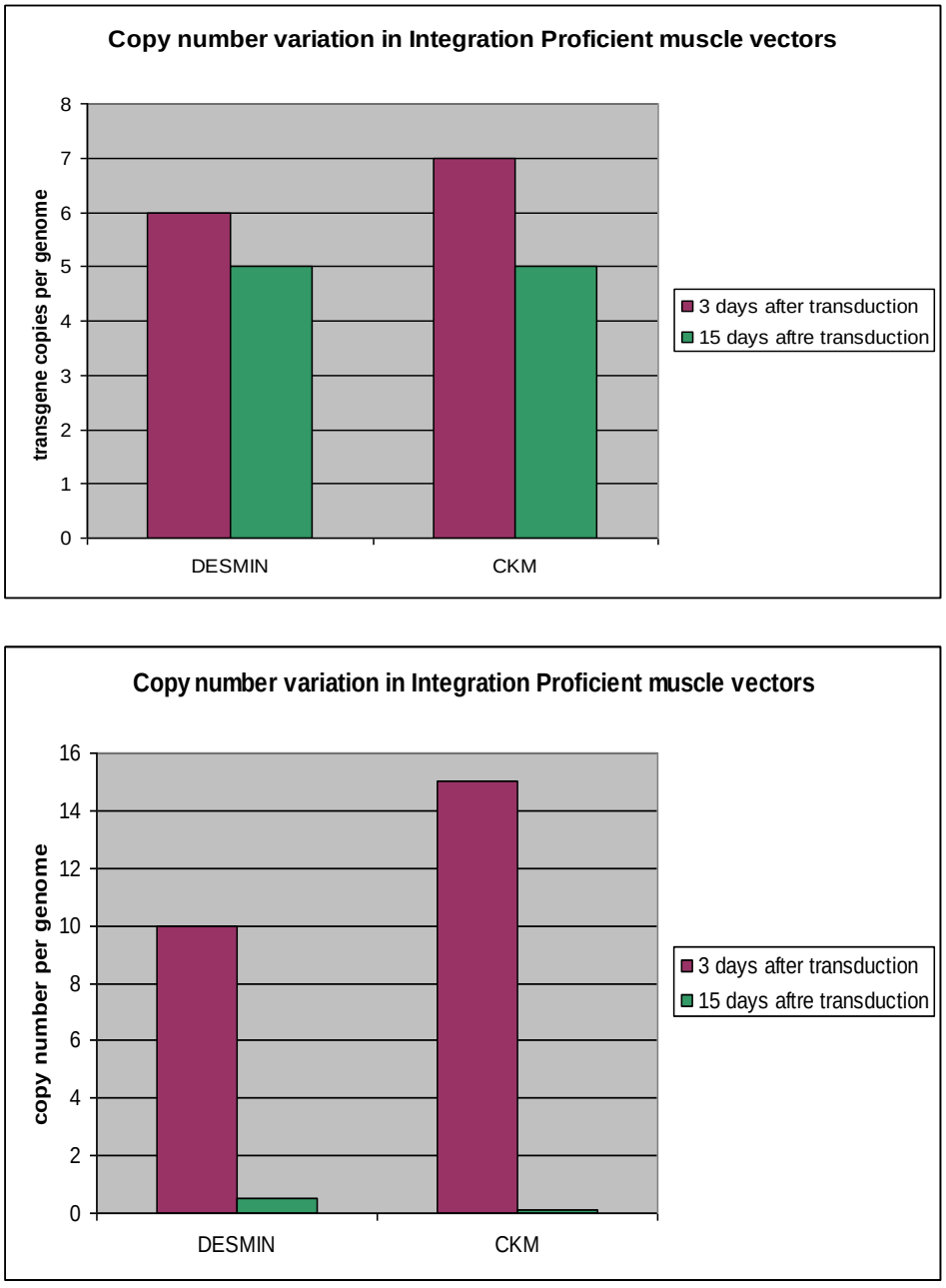


Figure 3.22 - Copy number variation in C_2C_{12} cell pools transduced with IP (upper panel) and ID (lower panel) vectors (n=3).

3.7 Conclusions

3.7.1 IP-LV

3.7.1.1 Comparison of *HBB* driven IP-LV and the inclusion of an *HHB* specific element into a ubiquitous expression promoter

It is interesting to note that despite the inclusion of the 3' half of *HBB* with a heterologous cDNA sequence into the GLOBE-GFP vector, the 3'GLOBE-GFP construct did not show any significant increase in expression over the sense orientation construct (GLOBE-GFP). The two vectors have similar profiles for GFP expression and MFI, with comparable copy numbers.

Interestingly, the LCR-SFFV vector shows a lower value for CV (**Figure 3.17**), better than both HBB-based vectors (3'GLOBE-GFP and GLOBE-GFP) and the SFFV promoter alone (SEW), with far less variation between individual samples (reduced error bars). These results show the specificity of the LCR element and show that the element dominantly interacts with a heterologous viral, which is able to largely negate interference nearby host gene elements at the site of vector integration, thereby enhancing erythroid expression. The obvious differences between the two SFFV driven constructs are caused by the presence of the LCR.

3.7.1.2 Comparison of muscle specific IP-LV

The muscle specific LVs tested in this study show similar with regards to eGFP positive cells and copy number. The 2 vectors differ only as the CK-M driven vector showed higher values of MFI in IP context, suggesting that the expression per copy number in this vector is higher than the one of the *DES* promoter.

3.7.2 ID-LV

All vectors used in this study either ubiquitously expressing (SFFV driven) or tissue specific (*HBB* and muscle *DES* / *CK-M* specific), were produced without major complication and expressed from both IP an ID configurations. As initially expected, the expression of ID constructs decreased significantly following successive rounds of cell division and reached values close to zero in some vector configurations.
Nevertheless, this study demonstrates that tissue specific as well as ubiquitously promoters can express from an ID context initially, even if at low intensity levels.

3.7.2.1 Vector ability to express in an integration deficient configuration can be extended to a tissue specific context

All constructs are able to express initially in an ID configuration. This expression is though, in general at a lower level per vector copy when compared to the same cassette packaged as a standard integration LV.
ID expression is variable; note, for instance, the size of the error bars in **Figure 3.15** (lower panel) and **Figure 3.19** (lower panel) indicating high variability between pool samples).

3.7.2.2 Comparison of ID-LV driven by muscle specific promoters

Overall both muscle constructs express when in an ID-LV configuration, at least at early time points following transduction, however this expression drops 3 days after transduction. To be sustained and harbour expression for a longer time period in an ID context, muscle driven vectors must transduce at higher MOI values (MOI of 1 used in this set of experiments).

3.7.2.3 *HBB* LV constructs *versus* muscle LV constructs in ID configuration

According to the results presented over a period of 15 days, the percentage of eGFP positive cells in *HBB* driven constructs drops more abruptly from day-3 following transduction (**Figure 3.15**) when compared with the muscle specific *DES* and *CK-M* vectors (**Figure 3.19**). Although, it is should be noted that the initial vector copy number for both muscle vectors in C_2C_{12} cells is 10 and 14 respectively for *DES* and *CK-M*, (**Figure 3.22**) which is higher than the values for obtained for the *HBB* driven constructs (**Figure 3.18**). Therefore, it will take longer for dilution of ID-LV through cell division in the case of the muscle than the *HBB* LVs.

It is therefore possible to say that according to our results, the higher the initial copy number values of a transduced cell population, the slower the drop in vector in the sample and thus the longer is the duration of expression from ID-LV.

3.7.3 Concluding Remarks

- Our experiments provide evidence that the utility of ID-LV can be extended to provide tissue-specific as well as varying degrees of ubiquitous therapeutic gene expression depending on the control elements used.

- Expression levels from all vectors analysed in an ID-LV configuration is always lower than the corresponding vector in an IP-LV configuration.

- Duration of expression from ID-LV within a pool of cells is dependent on initial MOI.

3.8 Consequences for potential use of ID-LV for gene therapy of the haemoglobinopathies

During this series of experiments the results obtained with non-integrating lentiviral vectors although positive, showed that expression from within an ID context is far lower per vector copy than with the same construct delivered through a standard integrating configuration.

Given these results is was clear that a high vector copy number per cell would have to be delivered to achieve a therapeutic effect for the haemoglobinopathies. This is in addition to assuming that the problem of ID-LV replication and retention can also be solved.

It was therefore decided that that the construction of a therapeutic globin LV in an integration deficient configuration would fail to provide therapeutic values for rescue of erythropoiesis and the project certainly needed to evolve into other research pathways.

With all this in mind, we decided to focus our attention on the improvement and efficient delivery of integration proficient *HBB* LV constructs.

Chapter Four

COMPARATIVE ANALYSIS OF UCOE-BASED LENTIVIRAL VECTORS IN A MURINE NEONATAL INTRAVASCULAR DELIVERY MODEL SYSTEM

4.1. Aims of Chapter Four

- Perform a comparative analysis of UCOE-based lentiviral vectors in a murine neonatal intravascular delivery model system

- Find ‘the ideal vector’ for further studies *in utero*.

4.2 Introduction

Ubiquitous chromatin opening elements (UCOEs) consist of methylation-free CpG islands extending over dual divergently transcribing promoters of housekeeping genes and have been found to be resistant to heterochromatin-mediated silencing (*Antoniou et al., 2003; William et al., 2005; Lindahl Allen & Antoniou, 2007*). Recent work has shown that the human *HNRPA2B1-CBX3* UCOE (A2UCOE) gives rise to reproducible and stable expression that is resistance to silencing from within a self-inactivating (SIN) lentiviral vector context (*Zhang et al., 2007; Zhang et al., 2010*).

In the search for the optimal vector to be employed for haematopoietic cell marking following *in utero* delivery, we conducted experiments via intravascular delivery in day-old neonatal mice to assess the reproducibility and stability of expression from a series of UCOE-containing lentiviral vectors either alone or linked to the human Cytomegalovirus (CMV) promoter/enhancer.

4.2.1 The UCOE element

The Ubiquitous Chromatin Opening Element (UCOE) used in these studies is an enhancer-less element that consists of a methylation-free CpG island encompassing dual divergently transcribed promoters of the *heterogeneous ribonucleooprotein HNRPA2B1 (A2) and chromobox homologoue 3 (CBX3)* housekeeping genes (*Antoniou et al., 2003; Williams et al., 2005*).

This UCOE is able to give constant and stable transgene expression in tissue culture cells (Antoniou et al., et al., 2003; Williams et al., 2005; Zhang et al., 2007 & 2010) and in mice (Katsantoni *et al.,* 2007; Zhang *et al.,* 2007 & 2010). In addition, it is also able to avoid transcriptional silencing and position effect variegation (PEV) expression pattern when integrated near or within centromeric heterochromatin

(Antoniou *et al.,* 2003). These observations imply that this UCOE functions as a dominant chromatin remodeling or opening element to prevent transgene silencing. However, only transgene constructs extending over both *HNRPA2B1* and *CBX3* promoters, and not *HNRPA2B1* alone, is able to give stable transgene expression and prevent from silencing (Antoniou *et al.,* 2003).

In 2003 *Antoniou* and colleagues (*Antoniou et al., 2003*) showed for the first time that transgenes containing a 2kb fragment spanning the dual divergently transcribed promoter region of the *HNRPA2B1-CBX3* (driving expression of a eGFP reporter gene) was sufficient to prevent transcriptional silencing and a variegated expression pattern when integrated into heterochromatin.

More recently it has been shown that the same 2.2kb *HNRPA2B1-CBX3* UCOE (A2UCOE) within an LV context produces a consistently high and stable expressing population of cells that were not prone to gene silencing *in vitro* and in HSC and peripheral blood cells *in vivo* (*Zhang et al.,. 2007; Zhang et al., 2010*). This is in marked contrast to LVs containing either SFFV or CMV-regulated constructs (eGFP transgene), which showed extensive silencing. These results suggest that this *HNRPA2B1-CBX3* UCOE (A2UCOE) was efficiently able to abolish progressive silencing and PEV of a stably integrated transgene over time in a cell line or *in vivo* and also reflects its dominant chromatin opening function. Furthermore, *ex vivo* delivery of an A2UCOE-IL2RG LV to HSC of a murine model of SCID-X1 was able to completely rescue the disease condition at low (1/cell) vector copy number (Zhang *et al.,* 2007).

Most recently, the A2UCOE has been shown to stabilise expression from the neutrophil-specific MRP8 promoter without compromising specificity and with an A2UCOE-MRP8-gp91phox LV able to efficiently rescue the disease phenotype in a mouse model of chronic granulomatous disease (Brendel *et al.,* 2011).

4.2.2 The Cytomegalovirus (*CMV*) promoter

The human cytomegalovirus promoter (*CMV*) is a ubiquitous promoter that contains a strong enhancer region and despite its vulnerability to position effects, is commonly used to drive reporter genes in various tissues and cells. It was therefore used as a constitutive comparison in this study. The construct used was a 7.4kb Sin-18 lentiviral vector pRRLsin ppt.*CMV*.EGFP (Follenzi *et al.*, 2000).

4.3 Vectors in this study

In this comparative study we used LVs previously constructed in the *Antoniou* group and/or by collaborator laboratories. We chose to use the A2UCOE given the reports of consistent and reproducible expression obtained with this element in HSC and peripheral blood cells *in vivo* ((*Zhang et al., 2007; Zhang et al., 2010; Brendel et al., 2011*). Vectors contained an eGFP reporter transgene driven either directly from the *HNRPA2B1* promoter (Antoniou *et al.,* 2003; Zhang *et al.,* 2007) or a CMV promoter with the core A2UCOE linked upstream of this element (Williams *et al.,* 2005) (**Figure 4.1**) .

The A2UCOE and CMV vectors were designed in the Antoniou laboratory, with the A2UCOE-EGFP and A2UCOE-CMV vectors provided by *Gillian Talbot* and the UCOE-EGFP-WPRE construct by *Pascal LeClere*.

The vectors employed are self-inactivating (SIN) due to a deletion of the U3 region within the 3' LTR. Therefore, only the internal promoter is active from the integrated provirus in transduced cells.

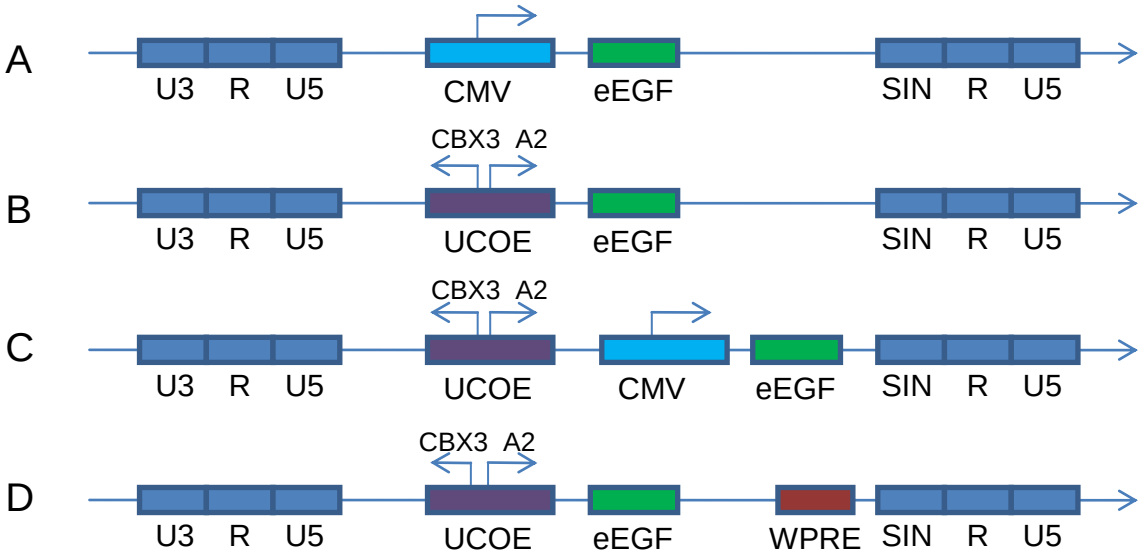


Figure 4.1 Diagrammatic illustration of lentiviral vectors. All vectors employ an enhanced green fluorescent protein (eGFP) reporter gene. **(A)** CMV promoter construct (Follenzi *et al.*, 2000) **(B)** A2UCOE driving expression from the *HNRPA2B1* promoter (Zhang *et al.*, 2007); **(C)** as (A) but with a 1.5kb core A2UCOE fragment linked upstream of the CMV promoter (Williams *et al.*, 2005); **(D)** as (B) but with the WPRE inserted downstream of eGFP. WPRE, woodchuck hepatitis virus post-transcriptional regulatory element. The vectors are self-inactivating (SIN) due to a deletion of the U3 region within the 3' LTR.

4.4 Results

Viral vector titer of EGFP-containing preparations was determined by transducing HEK293T cells cells with serial dilutions of virus and monitoring expression after 3 days FACS analysis. Titres ranged from between 5-10x10^8 transducing units/ml.

4.4.1 Vector studies performed *in vitro*

Lentiviral vectors were used to transduce HEK293T cells cells at a MOI to achieve 30-40% eGFP positive cells by FACS analysis. 3 pools of cells were transduced with identical amounts of virus per construct. Transduced cells were collected and FACS sorted at 10 days post-transduction.

4.4.1.1 *In vitro* Results

All vectors used in this study **(Figure 4.1)** were tested initially *in vitro*. **Figure 4.2** collates the images from the transduced wells and information obtained by FACS analysis, graphs with eGFP fluorescence and CV values.

Upon analysis of **Figure 4.2** it is important to note that the negative control sample shows a positive value for coefficient of variation (CV) due to a small number of cells being present in the positive quadrant in the histogram for eGFP expression. This is due to mis-placement of the gate arising from inexperience of FACS machine operation by the operator/student in training. One should have gated the negative population so that no cells were present in the positive quadrant and, only then, carry on with the FACS analysis. Nevertheless, this initial experiment is still informative as it gave an indication of relative vector performance.

The results show a clear difference between populations transduced with A2UCOE-based vectors and the CMV promoter alone construct and negative control. (Note: cells were transduced so that populations exhibit 20-40% eGFP positivity to avoid multiple copy integration per cell. Photographs were taken with the same exposure: t = 7s, f/4). This is a mere qualitative result but its evident that populations transduced with the A2UCOE-based vectors show an overall greater green fluorescence and are less prone to island clusters of EGFP positive cells (**Figure 4.2A**).

FACS plots showing eGFP expression of each transduced cell population can be found in **Figure 4.2B**. Again quantitatively it is possible to say that both CMV-UCOE and UCOE-WPRE show a much higher and consistent expression, with cell populations expressing equally. The cell population transduced with the A2UCOE-eGFP vector shows a lower but more consistent pattern of expression than that obtained with the CMV-eGFP vector as evidenced by the lower CV value (191 for CMV vs 123 for A2UCOE) and higher expression at lower emission levels and scattered from 10^1 to 10^2.

The CV values represent the variation amongst positively transduced cells, or the difference between individual cells in a given cell pool; that is, how much does each gated cell differ from another in relative level of expression. **Figure 4.2C** shows sample variation when transduced with different vectors - both constructs containing the CMV promoter (CMV and CMV-UCOE) display 2 individual peaks of cell populations whereas in wells transduced exclusively with the A2UCOE vectors (UCOE and UCOE-WPRE) show a single peak.

CV values alone also show the reproducibility in the UCOE and UCOE-WPRE vectors, with the CMV vectors presenting much higher variation (191.47 and 160.93) than the exclusive UCOE constructs (123.47 and 137.06).

These *in vitro* studies show that the addition of the WPRE has a positive effect on reporter gene expression. There is a clear shift to higher levels of expression when the WPRE element is included in the A2UCOE vector (**Figure 4.2 A, B, C**).

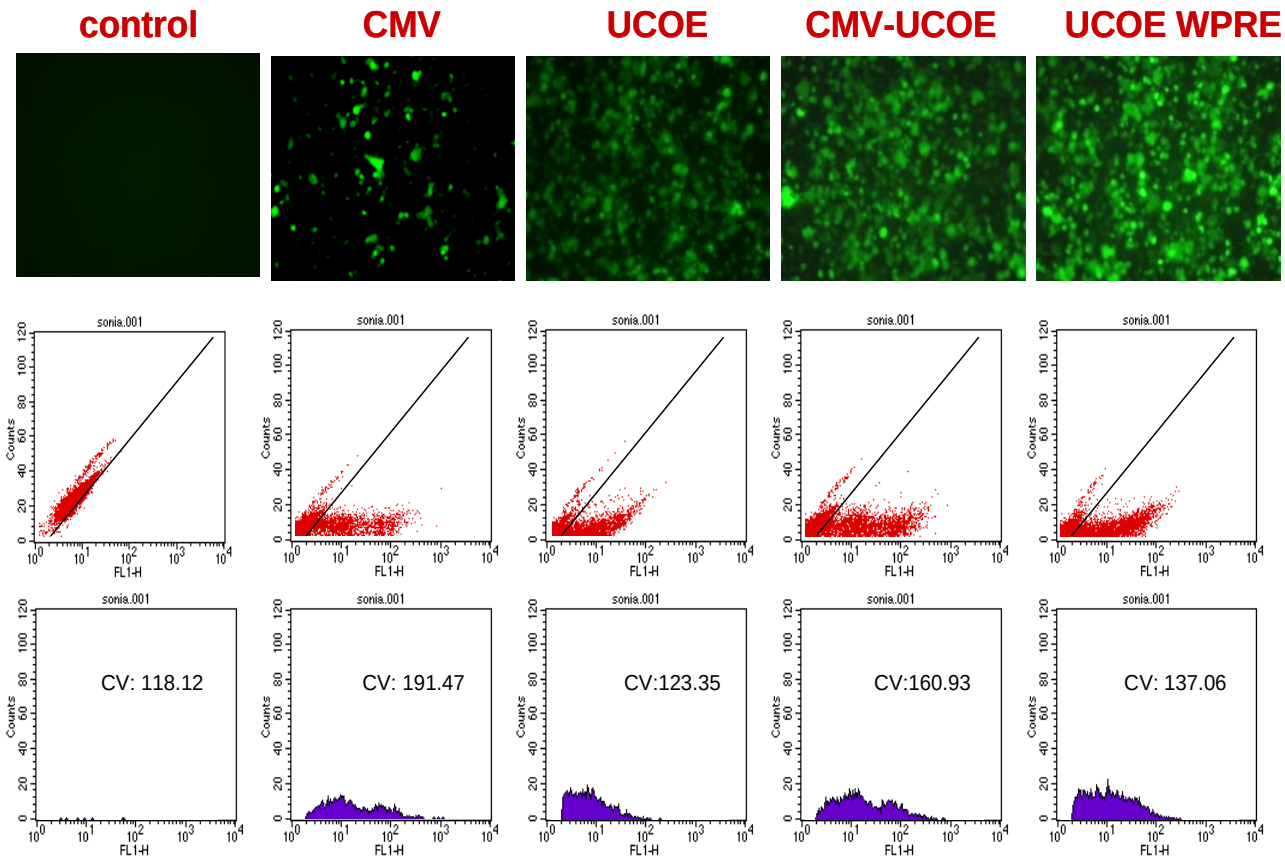

control
CMV
UCOE
CMV-UCOE
UCOE WPRE
sonia.001
Counts
FL1-H
CV: 118.12
CV: 191.47
CV:123.35
CV:160.93
CV: 137.06

Figure 4.2 Expression analysis of lentiviral vectors in HEK293T cells. HEK293T cells were transduced with the different lentiviral vectors (**Figure 4.1**) at a multiplicity of infection to achieve 30-40% eGFP positive cells by flow cytometry analysis. **(A)** Fluorescence microscope images taken with exposure: t = 7s, f/4. **(B)** Flow cytometry plots of percentage eGFP positive cells. **(C)** Flow cytometry analysis of tranduced cell pools showing values of coefficient of variation. Control meaning untransduced population.

4.4.2 Vector studies performed *in vivo*

The results obtained with the 4 vectors *in vitro* (**Figure 4.2**) were next expanded on by conducting a comparative study *in vivo*.

4.4.2.1 Neonatal injections

Wild-type neonate mice were injected intra-vascularly with 40μl of the viral suspension ($1\text{-}2x10^{8}$) via the superficial temporal vein 6-12 h post- birth. The pups were anaesthetised by being kept on ice for a few minutes, as due to their small size and limited resistance the pups cannot withstand any other form of anaesthesia. The injection was performed under a stereo microscope taking care not to inflict damage to any vital organ/ tissue that would result in tissue malfunction/death.
A small number of pups were rejected by their mums post-injection.

Injected mice were sacrificed at 6 weeks post-injection. By then all mice showed normal tissue morphology and there were no visible signs of vector toxicity. IIndividual tissues were collected from the carcases for determination of vector copy number and transgene expression.

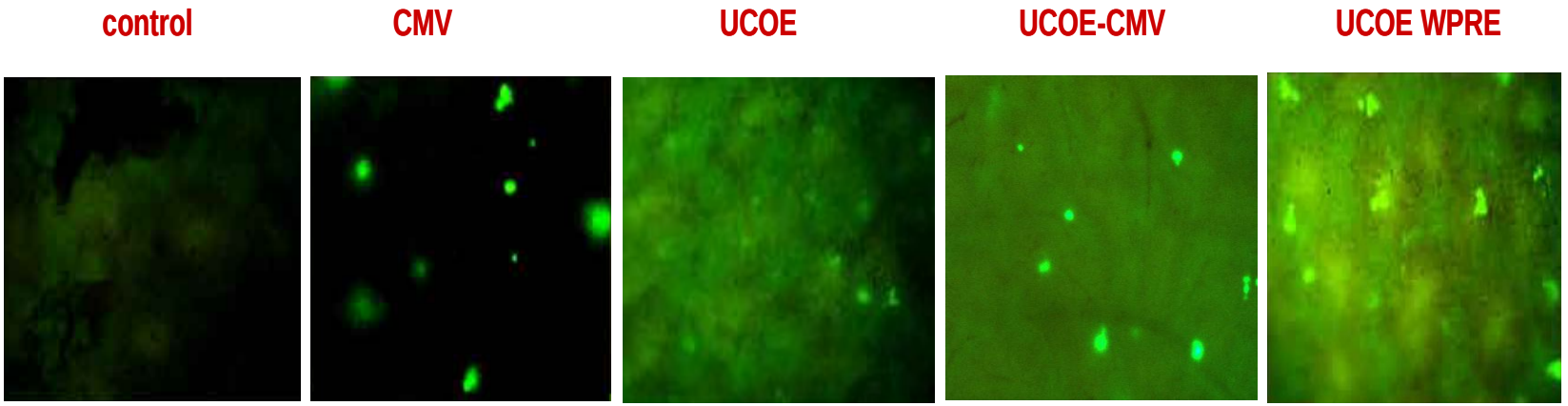


Figure 4.3 Fluorescence stereomicroscope images of liver samples from neonatally injected mice at 6 weeks post-injection. Wild-type mice were injected intra-vascularly via the superficial temporal vein with 40μl of viral vector suspension. Animals were sacrificed 6 weeks post-injection and freshly excised liver observed under a fluorescence stereomicroscope. Images are representative samples of each liver.

Figure 4.3 Fluorescence stereomicroscope images of liver samples from neonatally injected mice at 6 weeks post-injection. Wild-type mice were injected intra-vascularly via the superficial temporal vein with 40μl of viral vector suspension. Animals were sacrificed 6 weeks post-injection and freshly excised liver observed under a fluorescence stereomicroscope.

4.4.2.2 *Post-mortem* analysis (6 weeks after neonatal injection)

Liver, spleen, heart and lung were collected from animal carcases. Individual tissues were placed in PBS solution and macerated using a sonicator to obtain a smooth homogenate, from which DNA could be extracted (see section **2. 2.5.1**).
Vector copy number was determined by quantitative PCR employing the TaqMan system (see section **2.2.3.4.2** Chapter 2).

In order to determine eGFP protein levels, an ELISA approach was used and values normalised with respect to vector copy number.

4.4.2.3 eGFP expression per copy number

The results for copy number estimation via qPCR of both liver and spleen, are shown in **Tables 4.1** and **4.2**. Vector presence was not detected in either heart or lung tissues in the qPCR runs performed and neither was eGFP protein found in these organs by ELISA analysis.

Liver

CMV	UCOE	CMV UCOE	UCOE WPRE
1.3	1.3	0.7	1.5
0.7	1.2	0.6	0.8
1.3	0.7	1.2	0.7
2.0	-	1.5	1.2
-	-	-	0.9
Average 1.3	**Average 1**	**Average 1**	**Average 1**

Table 4.1 - Average LV copy number obtained by qPCR in liver tissues. Individual tissues were macerated and DNA extracted to determine vector copy number per cell using qPCR. Individual DNA samples were run in triplicate (n=3) and vector presence estimated by reference to an internal gene control (GAPDH) and a cell sample of known vector copy number value.

Spleen

CMV	UCOE	CMV UCOE	UCOE WPRE
0.7	0.6	0.6	0.3
0.3	0.3	0.6	0.7
0.3	0.6	0.5	0.7
0.5	-	1.1	1
-	-	-	0.6
Average 0.4	**Average 0.5**	**Average 0.7**	**Average 0.7**

Table 4.2 - Average LV copy number obtained by qPCR in spleen tissues. Individual tissues were macerated and DNA extracted to determine vector copy number per cell using qPCR. Individual DNA samples were run in triplicate (n=3) and vector presence estimated by reference to an internal gene control (GAPDH) and a sample of known vector copy number value.

It is known that injection via the superficial temporal results in fluid travelling into the mouse liver before spreading throughout the rest of the body via blood stream. Therefore, it is no surprise that large amounts of vector were found in mouse liver. Similarly, because the spleen is mainly responsible for the deposition of red blood cells, we were expecting large numbers of positively transduced cells in this organ.

Protein samples collected from mouse tissues were used to perform an eGFP ELISA to determine vector expression in individual organs. The results obtained for this experiment were normalized according to copy number and are shown in **Figure 4.4**.

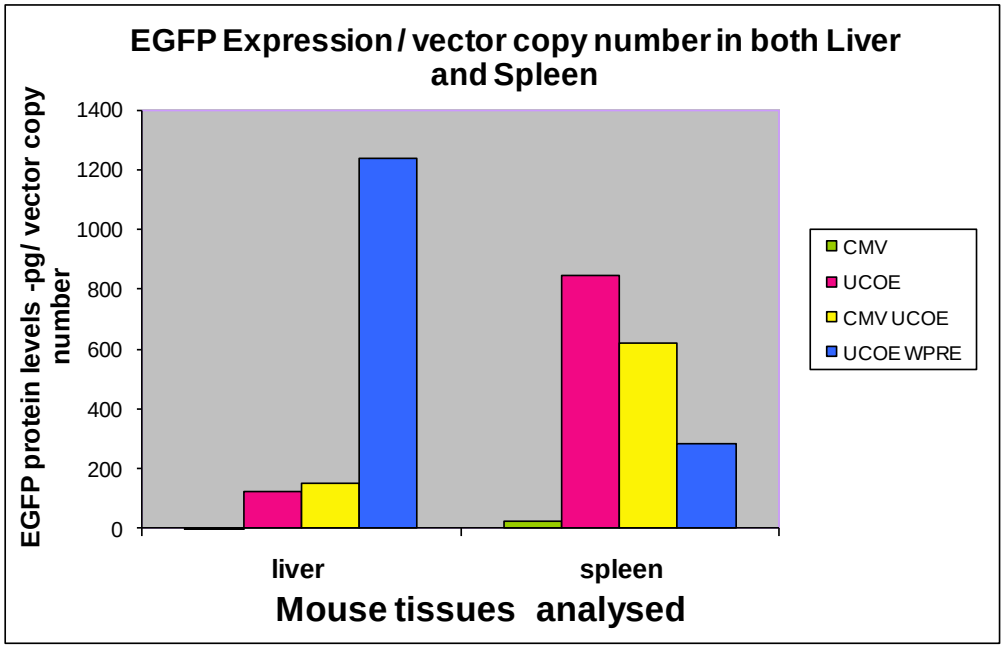
The CMV-eGFP construct showed relatively low levels of expression per LV copy number in both liver and spleen cells. Given that the viral vector preparations injected contained approximately similar amounts of viral particles in suspension, it is possible to say that this vector presents more difficulties in transducing murine cells than the construct in which the A2UCOE element is present.

UCOE-eGFP and UCOE-CMV-eGFP vectors performed similarly, with relative expression per copy in the liver and higher expression in the spleen. The UCOE-CMV vector showed expression levels very similar to the other UCOE vectors in this study, suggesting that the CMV promoter can be influenced by the UCOE element.

Finally, the A2UCOE construct containing the WPRE element (UCOE-WPRE) showed a surprisingly high eGFP expression per vector copy number in liver and a positive expression also in the spleen although lower than the other UCOE-containing constructs (**Figure 4.4**). One possible explanation to this result is that the transduced cells are still in the circulation (and somehow acquired advantage over untransduced cells in the animal's body). Unfortunately, we did not collect blood samples from these animals.

EGFP Expression (pg)/vector copy number		
	liver	spleen
CMV	2.2	25.5
UCOE	119.3	844.61
CMV UCOE	151.4	620.3
UCOE WPRE	1237.7	283.6

A



B

Figure 4.4 Expression of lentiviral vector constructs *in vivo*. DNA was extracted from murine tissues after systemic administration of lentiviruses and vector copy number per cell determined by qPCR employing the TaqMan system (**Tables 4.1** and **4.2**). eGFP protein levels were determined by ELISA and values normalised with respect to vector copy number. (**A**) The table presents the final values obtained and the graph (**B**) shows the eGFP expression per vector copy number. Vector presence and expression were not detected in either heart or lung tissue, which were also analysed.

4.5 Conclusion

4.5.1 Incorporation of A2UCOE avoids vector silencing

Given that the values for copy number for each vector were within the same range (**Tables 4.1** and **4.2**), the most probable cause for the lower values of expression per copy number with the CMV vector (**Figure 4.4**) is disturbance of the vector's expression by insertion site position effects. Indeed, previous studies have reported that this promoter is highly prone to DNA-methylation mediated silencing (Strathdee *et al.*, 2006: Zhang *et al.*, 2007*)*

All vectors driven by the A2UCOE element showed good eGFP expression in the 2 tissue samples analysed (**Figure 4.4**). The expression values vary according to different vector design but, in this experiment, expression is confined to the UCOE vectors. This observation together with previous studies (Nair *et al.*, 2011) showing promoter exclusivity suggest that the UCOE element is capable of rescuing the vector from eventual silencing caused by DNA methylation and insertion site position effects, blocking genes and enhancers *in trans*, preventing their interaction with the vector.

4.5.2 The search for an ideal vector for *in utero* injections

The ideal vector has to be capable of transducing effectively cells in the foetal liver since we are aiming to target HSC in this organ, before these cells migrate to the bone marrow compartment, which starts at approximately 14-16 days of gestation (refer to Introduction of both, Chapter 5 and 6).

Naturally hepatocytes are also going to be targeted by the LV in addition to HSC in the foetal liver. However, recent studies have shown that LV toxicity is very low and it's important to notice that all animals injected (n=3) presented normal tissue morphology, with no signs of tumours or malfunction.

4.5.1 Conclusion remarks

- Our data suggest that the A2UCOE is capable of rescuing insertion site position effects commonly associated with the CMV promoter, as well as on its own and to provide stable, consistent and reproducible levels of gene expression required for a safe, controlled gene therapy approach.

- A2UCOE driven LV construct containing both an eGFP reporter gene and the WPRE provided a highest degree of expression in comparison with the other vectors tested

We therefore decided to use the UCOE-EGFP-WPRE vector for *in utero* HSC marking experimentation.

Chapter Five

DEVELOPMENT OF IN UTERO GENE THERAPY APPROACHES

FOR INHERITED DISEASES – I

5.1. Aims

- Determine the optimal time of *in utero* gene therapy vector delivery in mice, prior to the switch of haematopoiesis from the foetal liver to bone marrow

- Test different reporter gene systems *in utero*

5. 2. Introduction

5.2.1 *In utero* gene therapy

Somatic gene delivery *in utero* is an approach to gene therapy for genetic diseases based on the hypothesis that prenatal intervention may avoid the development of severe manifestations of early-onset disease with permanent correction by stable transduction of relevant foetal progenitor cell populations.

In the investigation described in this Chapter, injections *in utero* aim to deliver LVs to the HSC population located in the foetal liver. This methodology takes advantage of the large scale migration of stem and progenitor cells to multiple organ compartments that occur during foetal development, with therapeutics providing treatment in time to prevent the development of severe disease manifestations. The approach is less likely to cause an immunological reaction due to foetal predisposition of immune tolerance to therapeutic transgenic proteins

5.2.2 Murine haematopoiesis

Delivery *in utero* aims to target the haematopoietic system at a specific site of haematopoiesis. Haematopoiesis is the process by which immature precursor cells develop into mature blood cells (**Figure 5.1)**.

Haematopoiesis in mice first occurs in the blood islands of the yolk sac and the aorta-gonad-mesonephros (AGM) region, 7 days post-conception (dpc) and lasts in these areas until 13 dpc. HSC activity is simultaneously detected in the placenta around 11 dpc. The next wave of haematopoiesis begins in the foetal liver at 12 dpc and in the spleen at15 dpc. HSC activity is subsequently detected in BM as early as 17.5 dpc, from where it is then sustained throughout the rest of life. The transition of haematopoiesis, migration

and relocation of HSC are thought to be regulated by chemokines as well as adhesion molecules (Kikuchi & Kondo, 2006).

In addition to the different sites of hematopoiesis, it is known that HSC acquire different properties and differentiation potentials during ontogeny; RBC derived from haematopoiesis in the yolk sac are primitive nucleated erythrocytes containing embryonic haemoglobin, whereas erythrocytes derived from the foetal liver or BM are non-nucleated cells containing only adult haemoglobin (Kikuchi & Kondo, 2006).

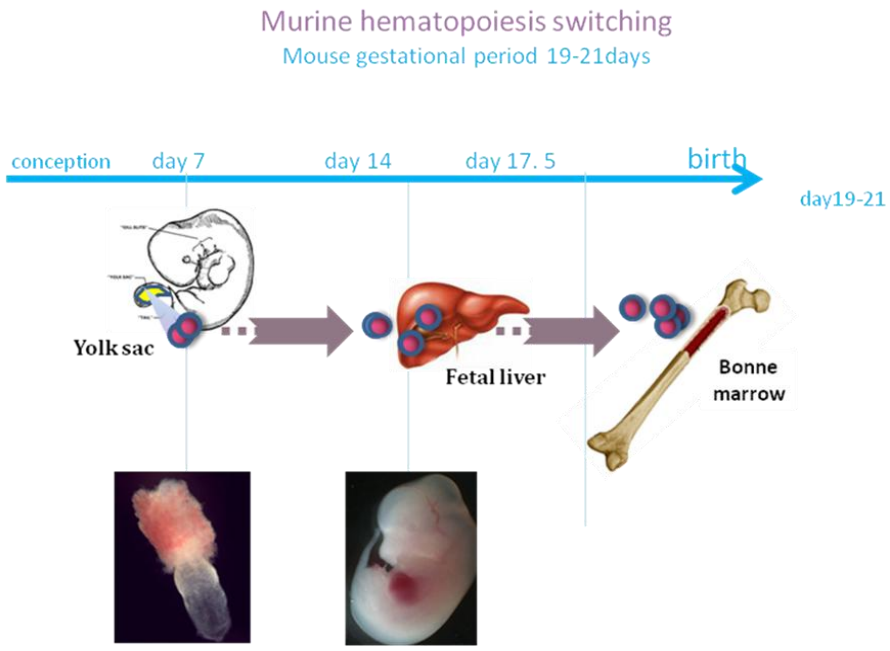


Figure 5.1. Schematic illustration of murine haematopoiesis switching. Mouse haematopoiesis starts in the yolk sac and shifts to the liver at gestational stage E14. A second switch in the site of haematopoiesis occurs again at E16-E17.5 with stem cells migrating to the BM.

5.2.3 Luciferase reporter gene

The aim of part of this project was to test A2UCOE-based vectors with different reporter gene systems, since each possesses distinct advantages. Luciferase was used as a reporter to assess *in vivo* transcriptional activity in mice injected *in utero* with LVs containing the luciferase gene driven directly by the *HNRPA2B1* promoter of the A2UCOE.

Luciferase emits light in the presence of a luciferin and ATP as substrate. Photon emission can be detected by a light sensitive apparatus such as a luminometer or modified optical microscopes. In this study imaging *in vivo* of treated mice employed the IVIS Imaging 50 Series (Xenogen) system, which measures photons from bioluminescence after intraperitoneal injection of luciferin. This system allows observation of biological processes *in vivo*, with the vector's presence and expression being observed while the animals are still alive in a totally non-invasive manner.

5.3 *In utero* injection of eGFP reporter gene vectors

5.3.1 LV used in this study

The initial *in utero* injections were conducted with the best candidate vector accessed previously using a neonatal delivery approach as described in Chapter 3. This SIN LV contains the A2UCOE driving expression of an eGFP reporter gene and a WPRE element as illustrated in **Figure 5.2**.

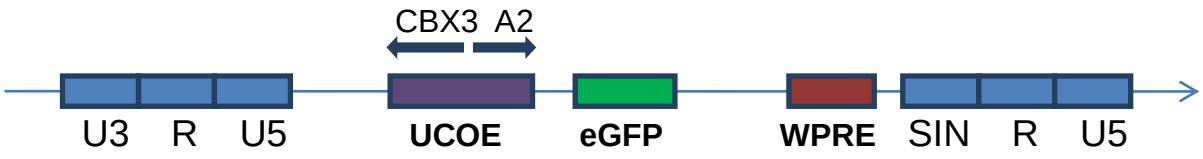


Figure 5.2 - Schematic illustration of the UCOE-eGFP-WPRE SIN LV. The vector contains the A2UCOE element driving expression of an eGFP reporter gene. The WPRE element is known to increase vector potency and transgene expression by facilitating transcript nuclear export.

5.3.2 Vector preparation

UCOE-GFP-WPRE was prepared as previously described (section **2.2.3**). Viral stocks were titred by serial dilution and concentrations determined by flow cytometry following transduction of HEK293T cells. Viral stocks generated ranged between 10^8-10^9 TU/mL.

5.3.3 *In utero* injections

Note: *In utero* procedures were conducted by Dr Simon Waddington.

Pregnant female MF1 mice at 14 and 16 days post-conception were used in this study. The procedure of intravascular injection into foetuses was similar to that previously described (Schachtner *et al*., 1999; Waddington et al., 2004; see section **2.2.4.2**, Chapter 2). Briefly, under isoflurane anesthesia, each horn of the gravid uterus was exposed through a full-depth midline laparotomy (**Figure 5.3**). A 34-gauge needle was used to perform a transuterine injection of 20μl viral suspension into the vitelline line peripheral yolk sac vessel.

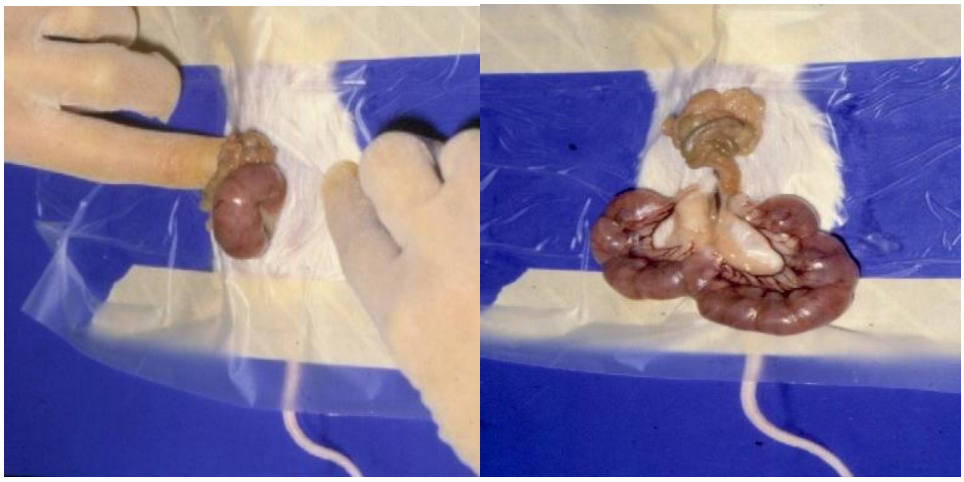


Figure 5.3. Exposure of murine uteri prior to foetal injection.

The laparotomy was closed in 2 stages, using interrupted stitches of 6-0 silk suture, and the pregnant females allowed recover in a warm cage.

5.3.4 *In utero* time points of delivery, E14 and E16

Initial *in utero* injections were a study to investigate targeting of the haematopoietic system. Various reports have been published describing the site of haematopoietic switching to and from the foetal liver (Kikuchi & Kondo, 2006; Yoder, 2007) but fail to time the migration of different HSC progenitors to and from foetal liver. An ideal timing of *in utero* injection should target not only the erythroid precursor cells but maximum numbers of HSC. We therefore decided to test 2 gestational time points, E14 and E16 (**Figure 5.4**).

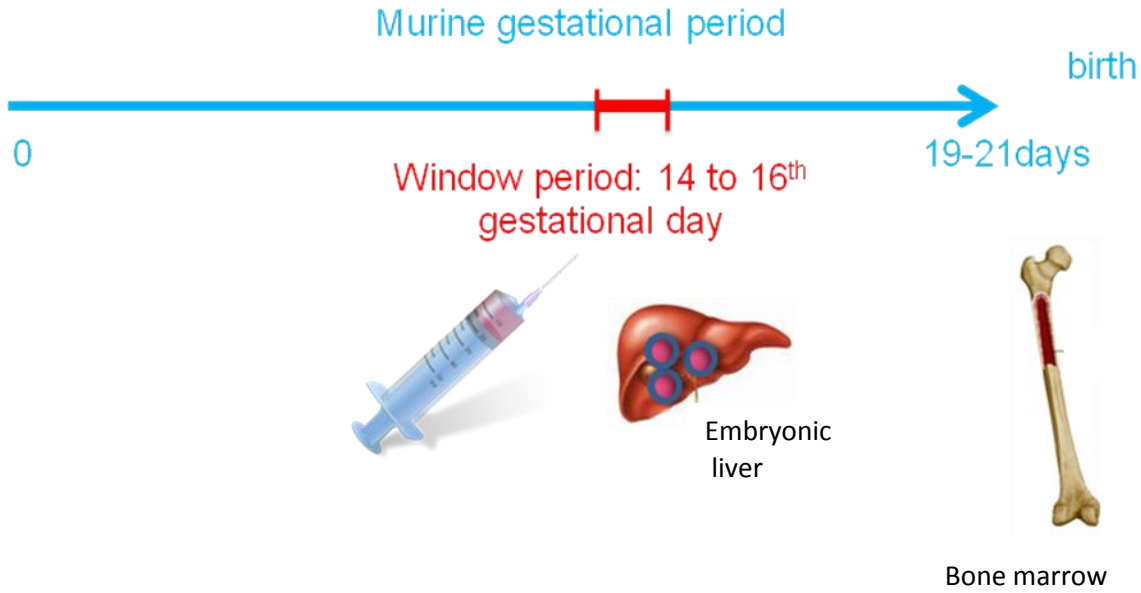


Figure 5.4 - Schematic illustration of *in utero* injections performed in pregnant mice. All *in utero* injections were performed during the time window of between E14 and E16, which is known to be when the foetal liver is the central organ of haematopoiesis.

5.3.5 Blood collection and analysis by flow cytometry

Blood was collected from mice 90 days post injection with animals under isoflurane anesthesia. Blood was mixed with anticoagulant solution (0.109 M sodium citrate buffer), mixed by vortexing and centrifuged. Samples were kept on ice until analysis.

Analysis of peripheral blood samples was conducted by staining with the appropriate antibodies for individual cell blood types, as shown in **Table 5.1.**

Antibody	Specificity
Ter119	Monoclonal antibody *Ter119* recognizing a component of glycophorin A and specifically marks late stages of the murine erythroid lineage (Kina *et al*., 2000)
CD11b	Monoclonal antibody *CD11b* recognizing mainly monocytes/ macrophages, granulocytes and microglia (Campanella *et al*., 2002)
CD19	CD19 protein is expressed on follicular dendritic cells and B cells. It is present on B cells from earliest recognizable B-lineage cells during development to B-cell blasts but is lost on maturation to plasma cells (Tedder & Isaacs, 1989)

Table 5.1 - Antibodies used in the staining of murine peripheral blood for flow cytometry analysis.

Collected blood was divided into 50μl aliquots for staining as described in section **2.2.5.7**, Chapter 2. Briefly, whole blood was used for *Ter119* staining. Staining for *CD11b* the cell aliquot was lysed prior to any antibody incubation. After washing and centrifugation, cells were incubated with Fc Block (to reduce background binding, and

increase murine specificity of the antibody). Samples were further incubated with control isotype antibody (coupled with *PE, APC* or *FITC* as appropriate) or specific antibody followed by another wash and fixation step in PBS solution containing 4% of formaldehyde. Individual stained samples were then analysed by flow cytometry against known control standards.

5.3.6 Results obtained from cellular staining of peripheral blood collected from mice injected *in utero* with the UCOE-eGFP-WPRE vector, 90 days post injection

The results obtained from the cellular staining of peripheral blood collected from mice injected at E14 are shown in **Figures 5.5 to 5.9**

- **Figure 5.5** shows the overall pattern obtained for each injected mouse compared to the uninjected controls. (The figure displays a representative mouse - *mouse number1*).

- The graphs in **Figures 5.6** to **5.9** depict the percentage of cells expressing positively for a given antibody - *Ter119*, *Cd11b* and *CD19* (the former exclusively at E16) - and eGFP simultaneously, and compared with the percentage of cells in the sample expressing eGFP alone.

5.3.6.1 Cellular staining of peripheral blood collected from mice injected *in utero* (E16) - representative sample

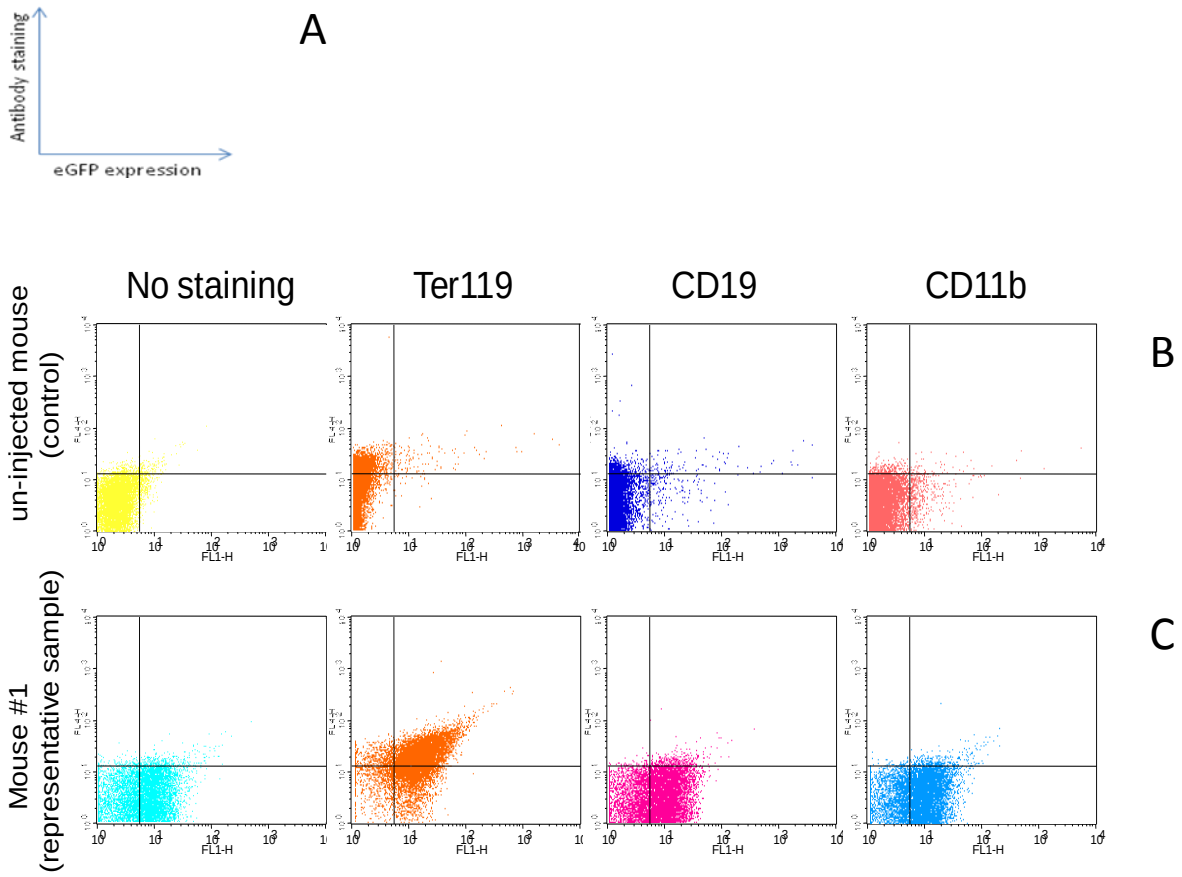


Figure 5.5 - Staining of peripheral blood cells collected from a representative mouse injected *in utero* at E16. A) The plots compare blood staining with different conjugated antibodies (FL2) against expression of eGFP (FL1). **B)** Upper panels represents the average staining pattern of an un-injected mouse, with the gate set with regards to the first panel (yellow) (injected, unstained) and variation from this staining pattern is shown in the remaining plots. **(C)** Lower panels *Ter119* antibody detects mature erythrocytes, *CD19* detects *B* cells and *CD11b* the monocyte population. FL1= eGFP ; FL2= APC or PE. As the results are similar for each stained peripheral blood sample, one representative profile is presented. Individual results for each mouse can be found in **Figures 5.6 - 5.9.**

5.3.6.2 Cellular staining of peripheral blood collected from mice injected *in utero* at E14 at 90 days post injection

5.3.6.2 .1 Cellular staining with *Ter119*

As can be seen in **Figure 5.6**, the percentage of cells with simultaneous *Ter119* staining and eGFP expression is significantly higher than the (unstained) population expressing only eGFP, in some cases reaching values close to 10-fold higher (mouse 1, 5, 6, 7 & 8). The graph compares two populations of cells; the population expressing only eGFP (unstained) and the population expressing eGFP and fluorescing positively for the presence of the erythroid precursor *Ter119*.

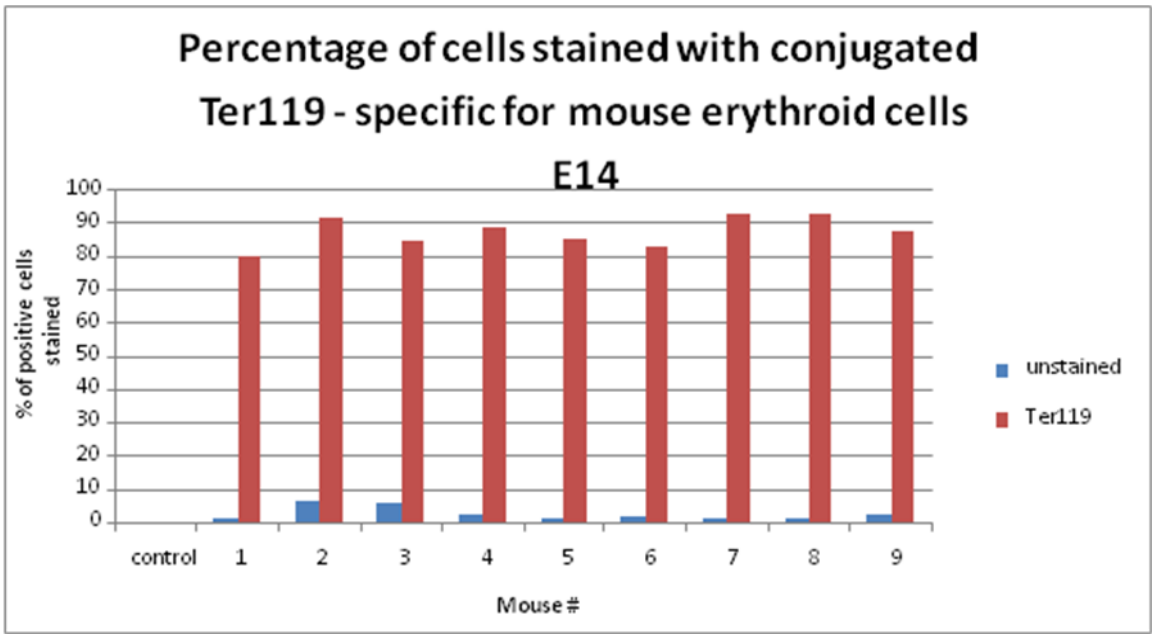


Figure 5.6 - Staining of peripheral blood cells collected from mice injected *in utero* at E14 with *Ter119* to highlight the erythroid population. Mice were injected *in utero* at E14 with 20μl of UCOE-eGFP-WPRE vector and subsequently bled at 90 days post injection. Control sample (blood from uninjected mice unstained) was used to gate the flow cytometry and used as a standard for the analysis of blood samples of injected mice, stained (red) and unstained (blue). The graph shows the variation of stained and unstained cell populations. Each pair of stained and unstained histogram bars depicts the result from an individual injected foetus.

5.3.6.2.2 Cellular staining with *CD11b*

CD11b staining of cells from E14 injected fetuses is shown n **Figure 5.7** The results show that the monocyte population in the blood sample is positively stained, in some cases up to 4-fold higher (mouse 1, 2, 6 & 7). However, mouse number 3 is an exception, which presented slightly more eGFP expressing cells than eGFP-*CD11b*.

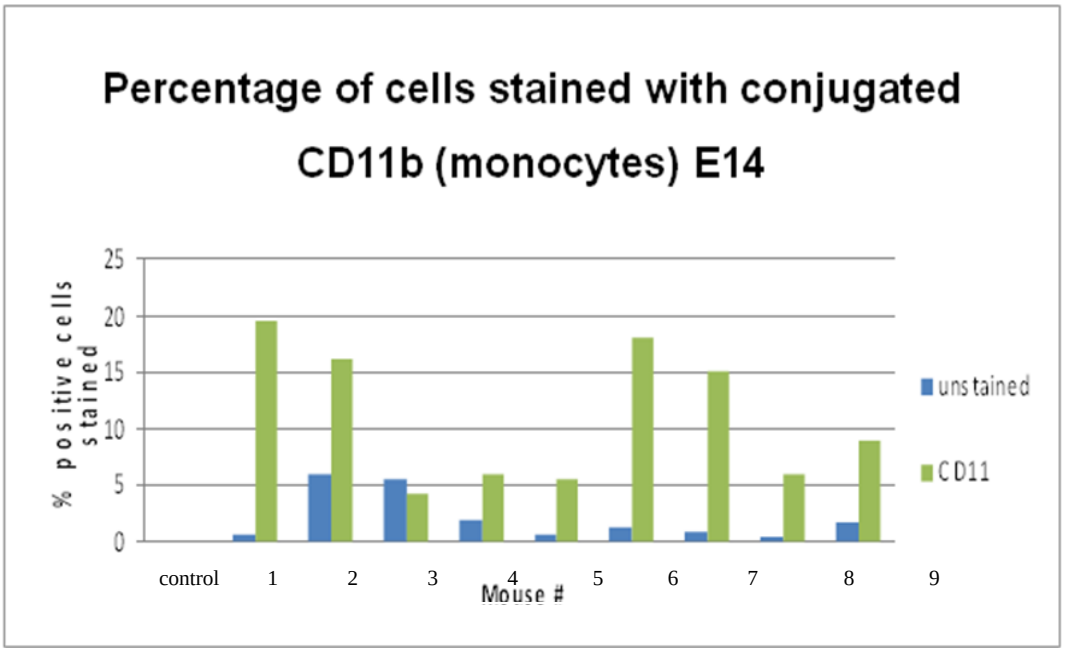


Figure 5.7 - Staining of peripheral blood cells collected from mice injected *in utero* at E14 with *CD11b* to highlight monocytes. Mice were injected *in utero* at E14 with 20μl of UCOE-eGFP-WPRE vector and blood samples collected 90 days post injection. A control blood sample from uninjected mice that was unstained for *CD11b* was used to appropriately gate the flow cytometry analysis and was used as a standard for the analysis of blood samples of injected mice, stained (green) and unstained (blue). The graph shows the variation of stained and unstained cell populations.

5.3.6.3 Staining of peripheral blood cells collected from mice injected *in utero* at E16 at 90 days post injection

5.3.6.3 .1 Staining of cells for *Ter119*

As before the number of cells positively staining for *Ter119* and simultaneously expressing eGFP is about 10-fold higher than the population expressing exclusively eGFP (**Figure 5.8).**

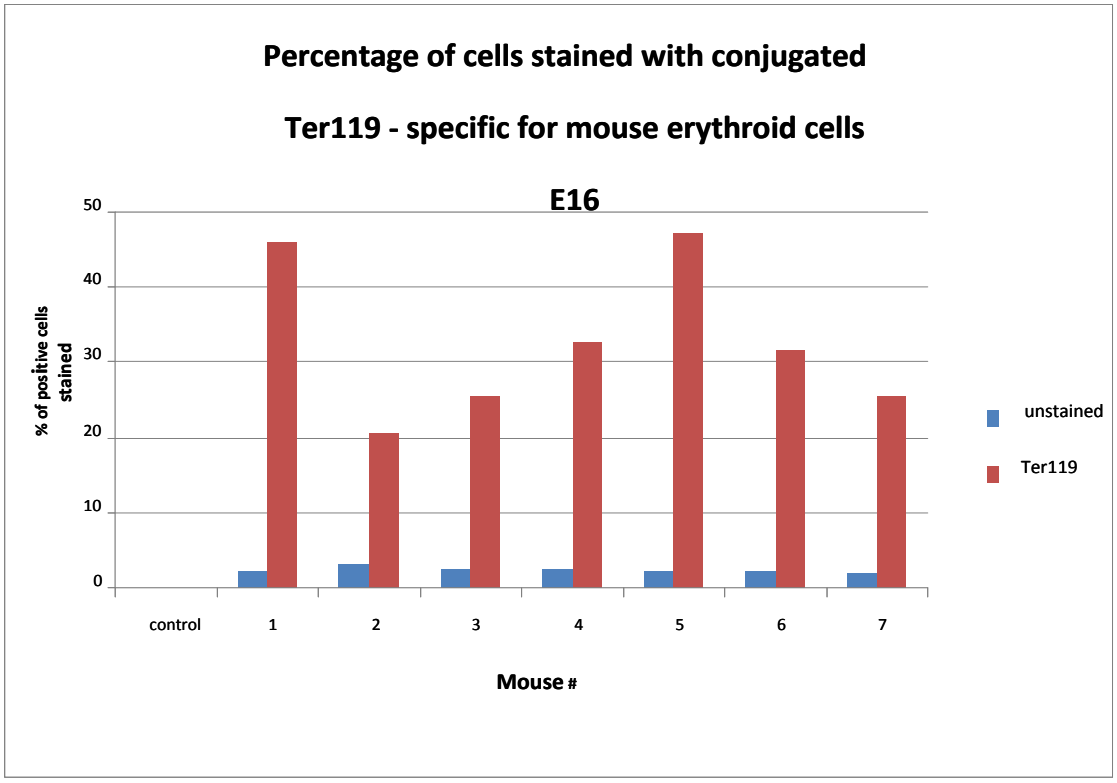


Figure 5.8 - Staining of peripheral blood cells collected from mice injected *in utero* at E16 for *Ter119* to highlight the erythroid population. Mice were injected *in utero* at E16 with 20µl of UCOE-eGFP-WPRE vector and bled 90 days post injection. A control blood sample from uninjected mice that was unstained for *Ter119* was used to appropriately gate the flow cytometry analysis and used as a standard for the analysis of blood samples of injected mice, stained (red) and unstained (blue). Each pair of histogram bars shows the result of stained and unstained cell populations.

5.3.6.3.2 Cellular staining for *CD19 and CD11b*

The number of cells positively staining for *CD19* and simultaneously expressing eGFP is overall 1-fold higher than the population expressing exclusively eGFP (**Figure 5.9)**. In contrast, the number of cells staining for *CD11b* and expressing eGFP is overall not significantly higher than the eGFP-only expressing population (unstained).

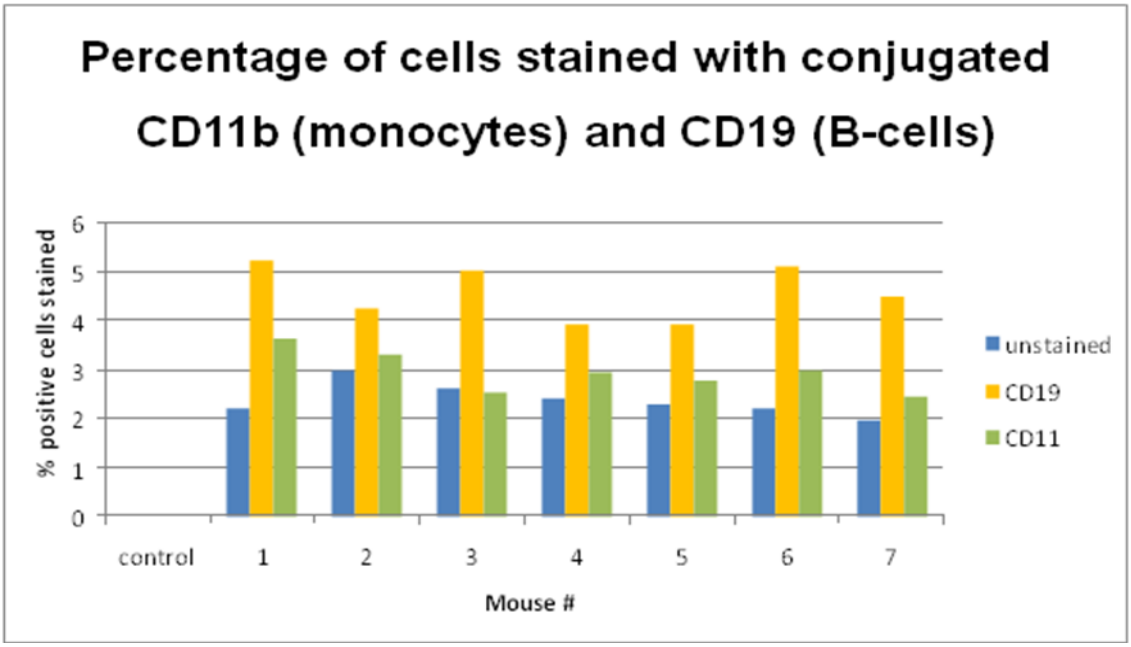


Figure 5.9 - Staining of peripheral blood cells collected from mice injected *in utero* at E16 with CD19 and *CD11b* to highlight B-cells and monocytes respectively.
A control blood sample from uninjected mice that was unstained for *Ter119* was used to appropriately gate the flow cytometry analysis and used as a standard for the analysis of blood samples of injected mice, stained with *CD19* (yellow) and *CD11b* (green) and unstained (blue). Each triplet of histogram bars shows the result of stained and unstained cell populations from an individual mouse injected *in utero* with 20µl of UCOE-eGFP-WPRE vector. Blood samples were collected 90 days post-natally.

5.3.7 - *Post-mortem* analysis 150 days after *in utero* injection

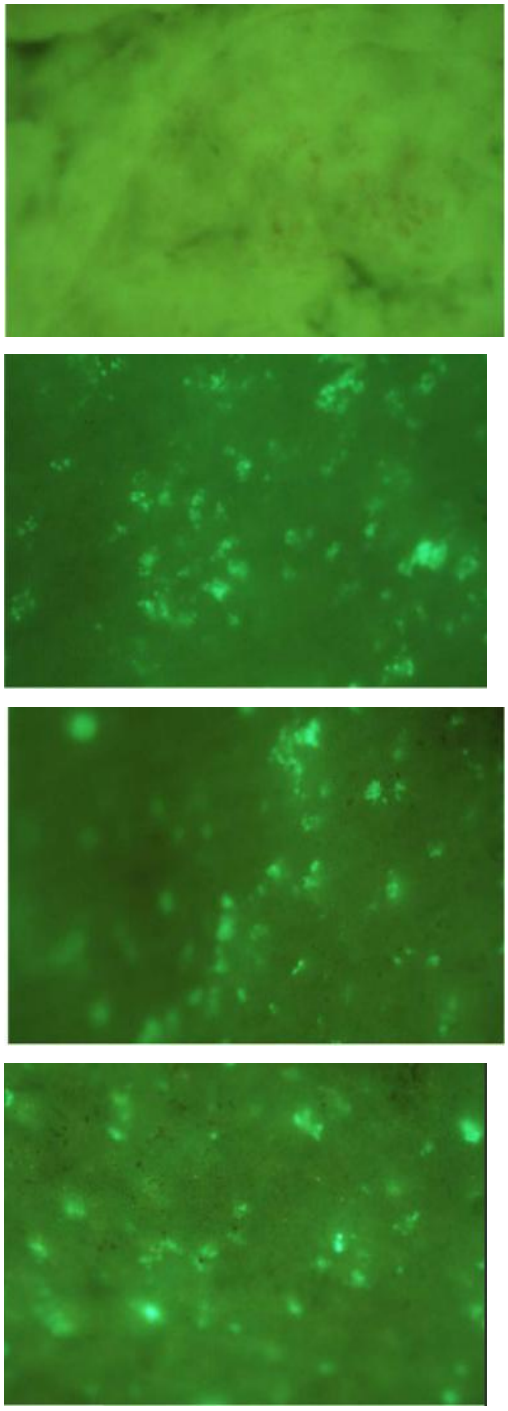
Animals were sacrificed by heart puncture so that the maximum amount of circulating blood was obtained. Liver, spleen, heart and lung tissues were collected from animal carcases. Individual tissues were submerged in PBS solution and macerated using a sonicator to obtain a smooth homogenate, without particulate suspension, from which DNA could be extracted.

Bone marrow cells were isolated from femurs and tibiae as described in section **2.2.4.7**, Chapter 2. Collected cells were used for proviral copy number estimation and stained for flow cytometry analysis.

5.3.7.1 Liver morphology

All mice showed normal tissue morphology. Livers collected *post mortem* were analysed by fluorescence microscopy to detect any eGFP positive cells. **Figure 5.10** shows representative images taken from the control mouse and injected test animals 1, 2 and 3.

The results show that injections performed *in utero* at gestational day E14 hit a higher percentage of precursor cells and therefore account for a higher number of eGFP positive cells in peripheral blood stained with *Ter119* and *Cd11b*.



Control (un-injected)

Mouse 1

Mouse 2

Mouse 3

Figure 5.10 - Representative pictures of livers expressing eGFP under fluorescence microscopy. Mice 1, 2 and 3 were used as representative samples since all livers showed a similar morphological pattern of expression. All tissues showed normal morphology in comparison with un-injected controls.

5.3.7. 2 Determination of proviral copy number

Mouse liver, spleen, heart and lung tissues were collected for viral vector copy number analysis using qPCR. Vector was detected in liver and spleen tissues as well in BM cells. The qPCR reaction was unable to detect vector presence in both lung and heart tissues. Copy number calculated and the results are shown in **Table 5.2**, below. Vector presence was not detected in either heart or lung tissues in the qPCR analyses performed.

Mouse	Liver	Spleen	BM	Lung	Heart
control	Undetected	Undetected	Undetected	Undetected	Undetected
1	2.582	0.528	0.895	Undetected	Undetected
2	2.145	0.125	1.32	Undetected	Undetected
3	2.761	0.236	0.481	Undetected	Undetected
4	2.021	0.514	0.624	Undetected	Undetected
5	1.527	0.951	0.472	Undetected	Undetected
6	2.158	0.751	1.593	Undetected	Undetected
7	1.925	0.467	1.593	Undetected	Undetected
8	2.352	0.524	0.587	Undetected	Undetected
9	2.470	0.825	1.020	Undetected	Undetected

Table 5.2. LV Copy number estimation in tissues of mice injected *in utero* at E14 with the UCOE-eGFP-WPRE vector. In the liver the number of LV copies per cell averaged 2.215 in the total injected population. In both spleen and BM the values obtained for proviral copy number are lower (average of 0.546 and 0.953, respectively for the whole population analysed). Vector presence was not detected in either lungs or heart.

Mouse	Liver	Spleen	BM	Lung	Heart
control	Undetected	Undetected	Undetected	Undetected	Undetected
1	2.945	0.852	0.524	Undetected	Undetected
2	2.594	0.423	0.210	Undetected	Undetected
3	1.894	1.085	0.631	Undetected	Undetected
4	2.396	0.892	0.127	Undetected	Undetected
5	2.127	0.954	0.784	Undetected	Undetected
6	1.545	1.237	1.214	Undetected	Undetected
7	2.395	1.295	0.958	Undetected	Undetected

Table 5.3. LV Copy number estimation in tissues of mice injected *in utero* at E16 with the UCOE-eGFP-WPRE vector. In the liver the number of LV copies per cell averaged 2.270 in the total injected population. In both spleen and BM the values obtained for proviral copy number were lower (average of 0.962 and 0.635, respectively for the whole population analysed). Vector presence was not detected in either lungs or heart.

These results clearly show that the injections were successful and the vector is present in mice tissues (liver and spleen) and BM cells. They prove that an *in utero* injection delivers vector onto these tissues, and, importantly for this project, efficiently transduces BM stem cells.

5.4 Comparing different reporter genes *in utero*: the UCOE-Luciferase-WPRE vector

The accurate, quantitative monitoring of expression of a given vector with an eGFP promoter is limited to isolated single cell populations such as peripheral blood and *post mortem* analysis. Following recent results obtained by collaborators we decided to incorporate in this study a vector that allows monitoring of its expression *in vivo* in which the eGFP reporter was replaced by a luciferase gene.

The 2.5kb A2UCOE (Antoniou *et al*., 2003) was obtained from the A2UCOE-eGFP vector by digestion with *Eco*RI and *Sal*I. Digestion of the plasmid SFFV with a multiple cloning site (SFFV-MCS-WPRE) with *Eco*RI and *Xho*I allowed removal of the SFFV promoter and creation of UCOE-MCS-WPRE. The 1.6kb luciferase cDNA was then amplified by PCR from SFFV-Luc-WPRE using primers that created *Sbf*I restriction sites at both the 5' and 3' ends of the PCR product. Digestion of the plasmid UCOE-MCS-WPRE with *Sbf*I allowed the ligation of luciferase and creation of the UCOE-Luc-WPRE construct (**Figure 5.11;** construct kindly provided by *Dr. Nathalie Ward*, ICH, UCL).

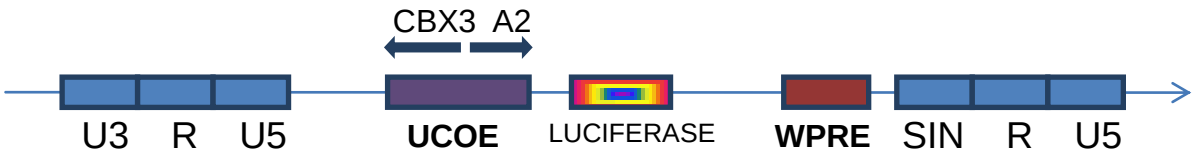


Figure 5.11 - Schematic illustration of the UCOE-Luciferase-WPRE SIN LV construct. The lentiviral vector construct contains a 2.5kb ubiquitous chromatin opening element (UCOE) from the human *CBX3-HNRPA2B1* locus acting as a promoter element driving expression a luciferase transgene (*vector courtesy of Dr. Nathalie Ward, ICH, UCL*).

5.4.1 UCOE-Luc-WPRE titer estimation

Preparations of the UCOE-Luc-WPRE LV were tested for titre by qPCR (section **2.2.3.4.2,** Chapter 2). This assay showed that vector stocks could be produced efficiently at an equivalent to 1.32 x 10^9 TU/mL.

5.4.2 *In utero* injections of UCOE-Luc-WPRE

The UCOE-Luc-WPRE LV was tested *in vivo* by *in utero* injection into foetal mice as described previously in this chapter for UCOE-eGFP-WPRE and in section **2.2.4.2** Chapter 2. Mice were injected with 20μL *Optimen* as negative controls.

5.4.3 Expression of Luciferase *in vivo*

Analysis of luciferase expression can be monitored over time using consecutive bioluminescent imaging. Quantification of *in vivo* luciferase expression was assessed using the IVIS Imaging 50 Series (Xenogen) system for bioluminescence detection after intraperitoneal injection of luciferin. Imaging was carried out for all animals at 30 and 360 days post-injection. Representative images from both time points are shown in **Figure 5.12** and **Figure 5.13**.

5.4.3.1 Results 30 days after *in utero* injection at E14 with the UCOE-Luc-WPRE lentiviral vector

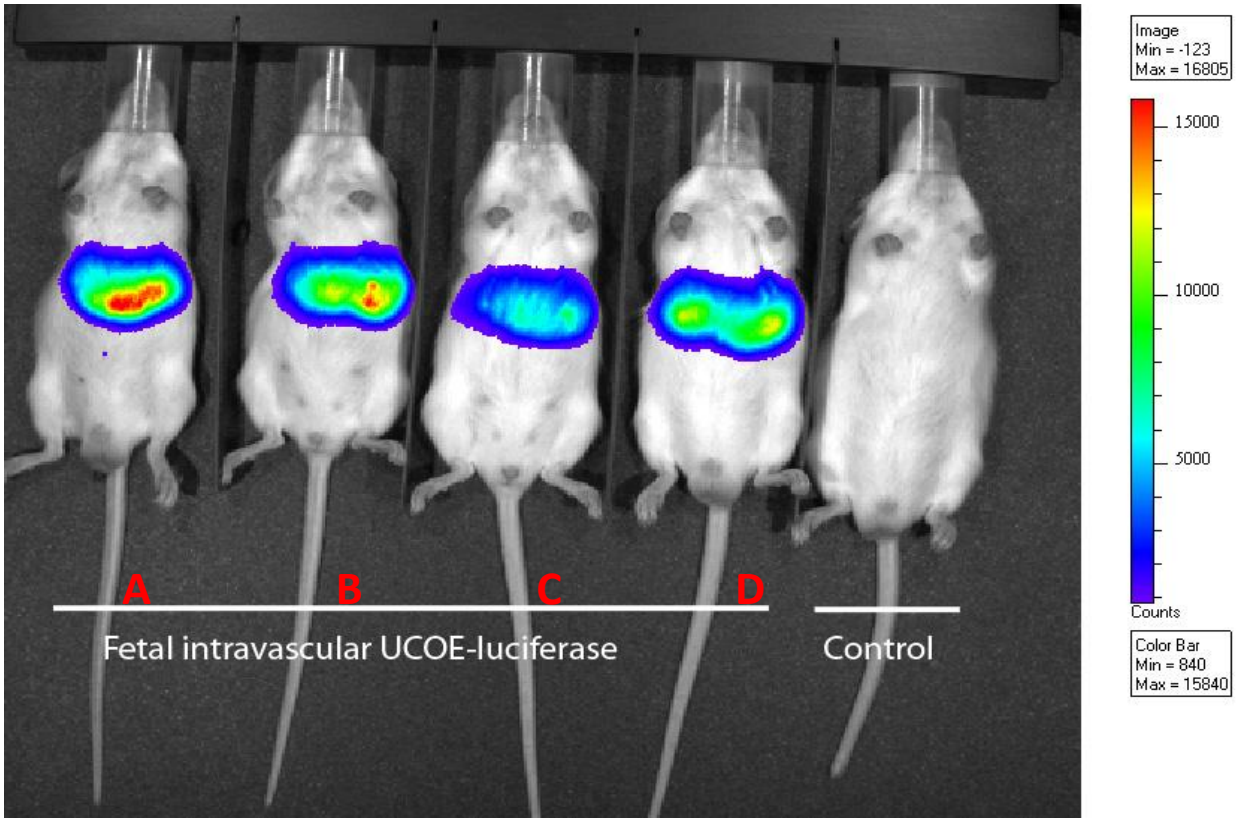
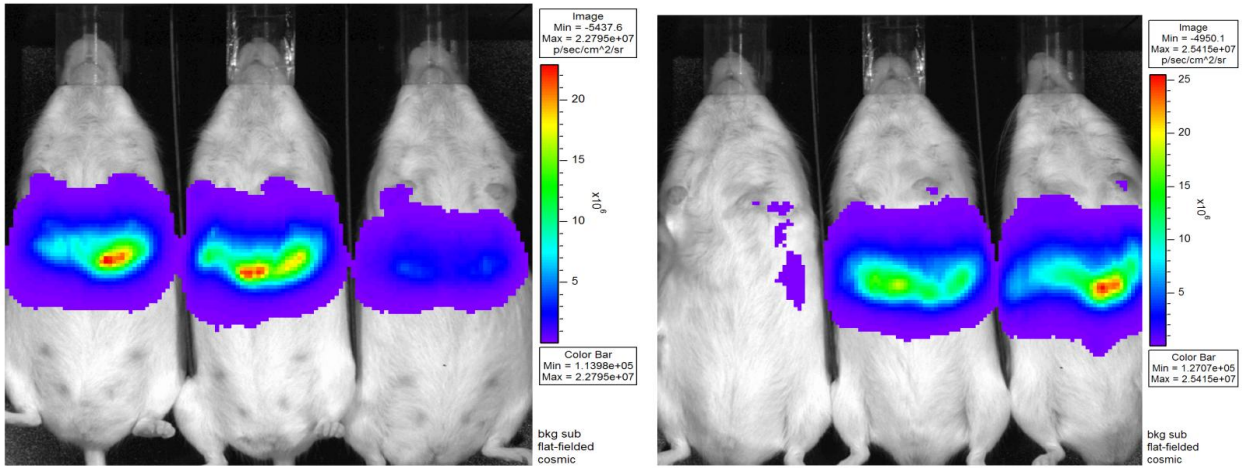


Figure 5.12 - Images of luciferase bioluminescence 30 days post *in utero* injection with UCOE-Luc-WPRE vector. Murine foetuses were injected at E14 and bioluminescence measurements taken at 30 days post-natally. Bioluminescence indicates the presence of actively expressed vector. A & D are male and B & C are female animals. The control mouse was injected with 20µL *Optimen*.

5.4.3. 2 Results 360 days after *in utero* injection at E14 with the UCOE-Luc-WPRE LV



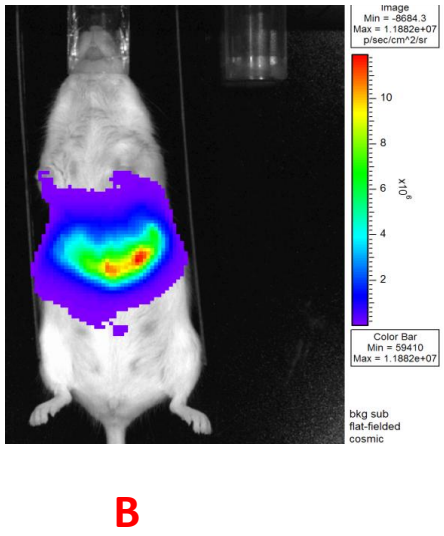


Figure 5.13 - Images of luciferase bioluminescence 360 days post *in utero* injection with the UCOE-Luc-WPRE vector. Mice A,B,C & D are those presented in **Figure 5.12.** Mice E & F were not subjected to bioluminescence measurements at day 150 post-injection. Bioluminescence indicates the presence of actively expressed vector. **Note:** the images presented have different color bars as the IVIS Imaging 50 Series (Xenogen) system automatically compares animals in a given shot and establishes aperture in accordance to expression. Image-brightness is therefore not comparable.

Figure 5.13 - Images of luciferase bioluminescence 360 days post *in utero* injection with the UCOE-Luc-WPRE vector. Mice A,B,C & D are those presented in **Figure 5.12.** Mice E & F were not subjected to bioluminescence measurements at day 150 post-injection. Bioluminescence indicates the presence of actively expressed vector. **Note:** the images presented have different color bars as the IVIS Imaging 50 Series (Xenogen) system automatically compares animals in a given shot and establishes aperture in accordance to expression. Image-brightness is therefore not comparable.

Luciferase protein activity per mouse 360 days after *in utero* injection is show in **Table 5.4**

Mouse	Average Radiance (p/s/cm^2/sr)
A	5.08 x10^6
B	2.47 x10^6
C	1.07 x10^6
D	4.6 x10^6
E	3.89 x10^6
F	4.54 x10^6

Table 5.4 - Luciferase activity values per mouse 360 days after *in utero* injection at E14. Luciferase values can be correlated with images in **Figure 5.13.**

All injected test mice had expression clearly well above background. *In utero* injection of the UCOE-Luc-WPRE vector in mice mediates expression in the central part of the body, most likely in the liver. No expression was observed in joints or BM. However, it is unknown if any expression was detectable in BM as the IVIS Imaging 50 system adjusts its analysis to the highest expressing areas and ignores/neutralises lower traces of expression.

5.4.4 *Post-mortem* analysis 360 days after *in utero* injection

Mice were sacrificed at 360 days post-injection and liver, spleen, heart and lung tissues together with BM cells were taken for analysis. Tissue samples were analysed for luciferase expression using the Luciferase Assay System (Promega) and a Bradford protein assay (BioRad) to give relative luciferase units (RLU) per µg protein. The number of integrated proviral copies was also assessed using quantitative real-time PCR (qPCR) to allow luciferase expression per LV copy to also be calculated.

5.4.5 Determination of integrated proviral copy number - luciferase injections

Liver, spleen, heart and lung tissues were taken from mice for determination of average viral vector copy number per cell using qPCR. **Table 5.5** shows the copy number for all tissues analyzed. Vector presence was not detected in either heart or lung in the qPCR analyses performed.

Mouse	gender	Liver	Spleen	BM	Lung	Heart
A	male	3.3645	1.532	2.0150	Undetected	Undetected
B	female	6.4815	1.384	0.0221	Undetected	Undetected
C	female	1.1582	0.128	1.9738	Undetected	Undetected
D	male	2.5696	0.743	3.4909	Undetected	Undetected
E	female	3.2704	0.341	0.8224	Undetected	Undetected
F	female	3.4739	0.524	1.6445	Undetected	Undetected
control	male	0.0911	Undetected	Undetected	Undetected	Undetected

Table 5.5 - LV average copy number per cell in tissues of mice following *in utero* injection at E14 with the UCOE-Luc-WPRE vector

5.4.6 Quantification of luciferase activity in tissues of mice injected *in utero* with the UCOE-Luc-WPRE vector

Luciferase reporter gene expression was determined as relative luciferase units (RLU)/μg total tissue protein and the results summarised in **Table 5.6.**

Tissue/organ compartment	Average RUL/μg protein
Liver	1.9×10^4
Spleen	4.8×10^3
BM	4.1×10^{3} *
Lung	Undetected
Heart	Undetected

Table 5.6 - Average relative luciferase units (RLU) per μg **total tissue protein.** The table shows the averaged values obtained for the whole set of tissues collected. Luciferase activity was not detected in either heart or lung tissues.***Note:** the average RLU value for BM included mouse B, which gave proviral copy number of 0.0221. If this sample is excluded the average RUL/μg protein would be 4.7×10^3.

RLU values obtained for each individual mouse are shown in bar histogram format in **Figure 5.14.**

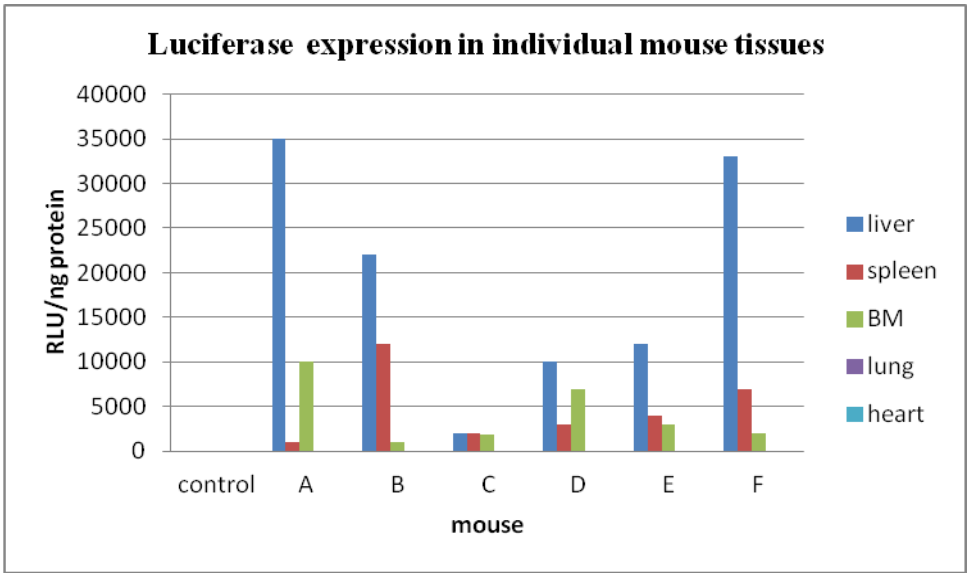


Figure 5.14 - Quantification of relative luciferase units (RLU) per μg total protein in individual tissues collected from mice injected *in utero* with the UCOE-Luc-WPRE vector. Liver, spleen, heart and lung tissues were collected from mice 360 days post-injection. Luciferase expression analysis using the Luciferase Assay System in conjunction with the Bradford protein assay with each sample analysed in triplicate.

Individual variation can be accessed by comparison of luciferase RLU units with values obtained for proviral copy number (**Table 5.5)**

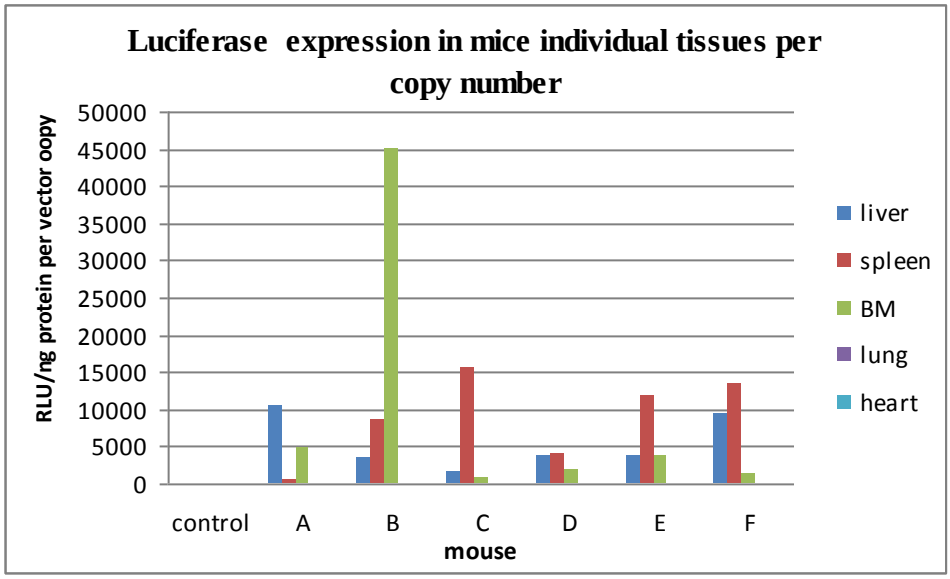


Figure 5.15. Luciferase protein expression per LV copy number in individual tissues collected from mice injected *in utero* with UCOE-Luc-WPRE. Mouse liver, spleen, heart and lung tissues were collected from injected mice 360 days post-injection. The values obtained for luciferase expression (**Figure 5.14**) were plotted against the values obtained for proviral copy number (**Table 5.4**) to give luciferase RLU/ng values per vector copy.

These results clearly show the presence of the Luciferase vector in mice tissues (liver and spleen) and vector engraftment into the BM tissues. Again, these positive results verify the feasibility of *in utero* delivery to target the hematopoietic population.

5.5 Conclusions

5.5.1 Injections with UCOE-eGFP-WPRE *in utero*

5.5.1.1 Liver morphology

The pictures obtained by fluorescence microscopy of mouse livers after *in utero* injection of UCOE-eGFP-WPRE (**Fig**ure **5.10**) differ significantly from those obtained by the administration of the very same vector to mouse neonates, where a homogeneous background for this vector was observed. (**Figure 4.3**, Chapter 4).

After *in utero* delivery, the UCOE-eGFP-WPRE vector expresses as isolated islands of what appear as colonies of transduced cells.

The titre of both vector preparations was similar so it is possible to say that the discrepancy between these two routes of delivery comes from the techniques used (*neonatal* versus *in utero* injections).

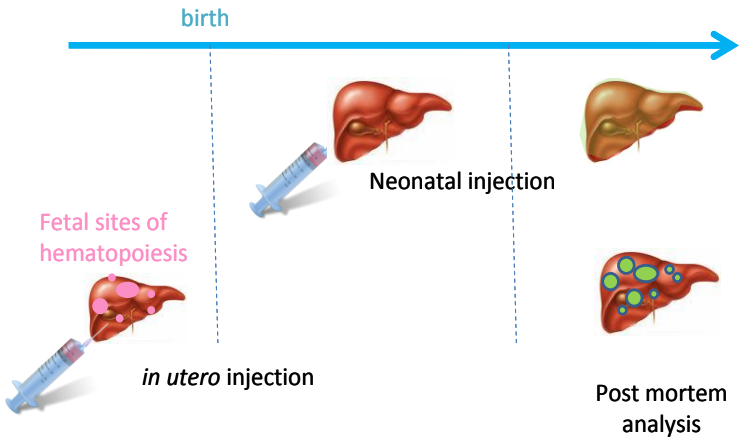


Figure 5.16 - Differences in UCOE-eGFP-WPRE vector expression patterns in liver after neonatal or *in utero* injection.

Differences in the cellular status of the liver between foetal and post-natal stages of development may contribute to the observed difference in UCOE-eGFP-WPRE vector expression patterns. The foetal liver is a mixture of both hepatocytes and haematopoietic stem and progenitor cells. Therefore, the more efficient transduction of hepatocytes over haematopoietic cells would give a distinct island, colony appearance, especially if this resulted from the expansion of a single transduced cell.

5.5.1.2 Staining of peripheral blood

Both the E14 and E16 time points of *in utero* vector delivery showed that a large percentage of cells in the circulation stain for *Ter119*, with a yet greater proportion of cells (10-fold higher) simultaneously staining for *Ter119* and expressing the eGFP reporter gene compared to eGFP alone (unstained). This result suggests that the great majority of mature red cells in the circulation are derived from HSC transduced whilst still in the foetal liver. Both time points showed positive results but at E14 the marking seems more evident with some mice showing dual *Ter119* staining and eGFP expressing over 90% of the cell population (mice 7 and 8).

For *Cd11b*, the staining pattern differed significantly between the two time points of vector delivery. While at E14 the difference between the stained and unstained populations was markedly different, at E16 the populations showed similar values. It is possible that monocyte precursors have started their migration to the BM by 16^{th} dpc and are therefore no longer present when the E16 injection took place. It is worth noting that switching of the site of hematopoiesis is not an immediate process but a gradual migration of cells and at a given point of murine gestation it is possible to find HSC and progenitors in more than one location.

Positive results shown by *Cd19* staining following injection at E16 confirm the feasibility of this hypothesis; that is, the injections *in utero* are transducing HSC or early

progenitor cell populations resulting in vector presence in more than one circulating cell type.

The values obtained for proviral copy number in different organ compartments can be found in **Tables 5.2** and **5.3**. The organ showing higher values of copy number in both E14 and E16 time points is the liver (averaged of 2.215 and 2.27, respectively). These values correlate well with the bioluminescence images shown in **Figure 5.10**, where liver tissues show evident expression of the eGFP reporter gene. In both spleen and BM the values obtained for proviral copy number are relatively lower. In spleen the average was 0.546 and 0.962 at E14 and E16 respectively and in BM average copy number was 0.953 and 0.635 respectively at E14 and E16. While vector detection in spleen confirms vector presence in the circulation, vector presence in the BM shows positive engraftment of the BM cells with the vector upon *in utero* injection, with a higher engraftment at E14. These results validate the hypothesis in this study.

Unsurprisingly, vector presence was not detected in either lungs or heart.

5.5.2 Injections with UCOE-Luc-WPRE *in utero*

5.5.2.1 Bioluminescence xenograph images

We decided to use a luciferase reporter gene LV as this allows *in vivo* documentation of vector expression in living animals. Animals were injected *in utero* with the UCOE-Luc-WPRE construct at E14 and vector expression analysed *in vivo* at both 150 and 360 days after injection (**Figures 5.12** and **5.13).** Both images show consistent expression *in vivo* and, more importantly, they show vector persistence 1 year after injection.

5.5.2.2 Proviral copy number

With regards to copy number the values obtained in the liver are not entirely a surprise, given the luciferase activity that is detected in the central part of each animal (**Figures 5.12** and **5.13).** The copy number values obtained in splenocytes are minimal and therefore cannot account for the bioluminescence visible in **Figures 5.12 and 5.13.** LV copy number values found in BM cells at 360 days post-injection are highly positive and validate the hypothesis behind these experiments; that is, vector delivery transduces HSC in foetal liver, which at a later gestational stage migrate to BM niches. Vector presence after such an extended (12 month) period of time also indicates that transduction is stable and not lost upon cellular division.

Interestingly, mouse B shows a high value of transduction in the liver and spleen but a low value in BM suggesting that vector transduction of HSC in the foetal liver has not occurred in his specific case.

5.5.2.3 Luciferase assay

Luciferase activity results in *post-mortem* tissues obtained from mice injected *in utero* with the UCOE-Luc-WPRE vector are very encouraging. As predicted by the *in vivo* images (**Figures 5.12** and **5.13**), luciferase expression in the liver is significantly higher than in any other organ compartments analysed.

Bone marrow cells transduced with UCOE-Luc-WPRE are clearly present as evidenced by the positive values obtained with the luciferase assay with this tissue.

Although luciferase activity in spleen cells is lower than in other organs, it is nevertheless indicative of transduced cells in the mouse blood circulation. Note that even though the bioluminescence image (**Figure 5.12 and 5.13**) suggests transduction of central body organs, tissues like heart and lung show neither presence nor expression from this vector. This result suggests that transduction of cells and tissues by *in utero* vector delivery via the route employed here, is confined to specific organs, and nearby tissues.

It is also interesting to note the lower RLU value in spleen cells in mouse A compared to other mice. For example, mouse B (RLU 1.2×10^4) luciferase expression is 11 times higher than that obtained in mouse A (RLU 1×10^3), even though the amount of vector injected was the same in both animals.

The evident lower level of transduction in mice can be explained by faulty injections of the vector; note that these injections are a difficult procedure on relatively small animals, which can account for the variation between different animals.

5.5.2.3 Luciferase expression units per copy number

Since the samples analysed are homogenates from the same tissue, and to account for individual variations, we plotted the values obtained for luciferase expression (**Figure 5.14**) against the values obtained for proviral copy number (**Table 5.5**) to obtain RLU values per vector copy (**Figure 5.15**).

Liver expression is constant in all animal samples analysed (blue bars), with the results presented not showing evident individual differences. The sample from mouse C shows lower vector expression in liver tissues. With the exception of mouse A, transduced cells in the spleen show expression values ranging from 5000-15000. These values indicate that although a relatively low percentage of cells in spleen are transduced (refer to the proviral copy number values, **Table 5.4**) their individual expression is high. Further integrational studies to determine possible vector integration sites and nearby transcription sequences may be informative with regards to the upregulation of the transgene in this case.

Expression per copy number in BM cells is represented by the green bars in **Figure 5.15**. With the clear exception of mouse B, BM cells show a lower expression per copy number within the range of tissues analysed. Nevertheless, these results are very encouraging as this degree of LV marking demonstrates the feasibility of an *in utero* delivery approach for targeting HSC to bring about early correction of disease. Mouse B shows an atypical high level of luciferase expression per LV copy number. Again, this sample would benefit from further vector integration analysis to access trasngene insertion sites and more importantly, if the pools analyses derive from different progenitor HSC cells.

5.5.3 Conclusion Remarks

- The use of eGFP vectors *in utero* has shown that both at E14 and E16 time points vector expression does not suffer significant changes relative to the amount of HSC and erythroid precursors transduced.

- On the contrary, cellular staining of the myeloid population indicates that at the E16 time point the population of transduced cells is significantly lower than at E14.

- The use of luciferase vectors allowed monitoring of vector presence and expression *in vivo*. We observed clear vector expression from the liver, and this visual result was later confirmed by both qPCR and luciferase enzyme assays.

- Both sets of experiments clearly document vector presence and expression in BM, proving that the starting experimental hypothesis was correct; that is, HSC can be transduced whilst in the foetal liver and then migrate to BM.

Chapter Six

DEVELOPMENT OF *IN UTERO* GENE THERAPY APPROACHES FOR INHERITED DISEASES - II

6.1 Aims

- Inject wild type (*MF1*) mice *in utero* with an LV containing an *HBB* expression cassette

- Analyse peripheral blood 150 days post-injection

- Sacrifice mice 250 days post-injection. Analyse peripheral blood, bone marrow and other relevant hematopoietic tissues.

6.2 Introduction

6.2.1 Previous studies with *HBB*-based LV - GLOBE

Previous studies with GLOBE LV performed in adult animals via an *ex vivo* bone marrow (BM) transplantation procedure have shown transcription levels of *HBB* sufficient to rescue severe β-thalassaemia *intermedia* and major phenotypes in mouse models of this disease (Miccio *et al*., 2008). In addition, studies have shown that successful transduction of GLOBE into $CD34^+$ cells from patients is capable of correcting the globin chain imbalance and restoring erythropoiesis upon differentiation (Roselli *et al*., 2010).

The success obtained with GLOBE *in vivo* led us to hypothesise that the administration of this or a similar vector by *in utero* injection via the vitelline placental vessel, would successfully transduce the HSC population located in the liver during foetal development, which are known to then migrate into BM niches around the time of birth (see **Figure 5.1**, Chapter 5).

The presence of the *HBB* transgene in BM, even if at a relatively low level, would theoretically mimic the positive results obtained previously by an *ex vivo* procedure (Miccio *et al*., 2008) and should therefore greatly reduce phenotypic effects of haemoglobin imbalance in the newborn.

6.2.2 Transcription termination

The percentage of human RNA polymerases capable of initiating and successfully completing gene transcription has been estimated to be in the order of 1% (Banerjee *et al.*, 2009). RNA polymerase II transcription complex activity is thought to be affected by chromatin modifying factors and intrinsic properties of the template and transcript that affect transcription elongation and termination (Banerjee *et al.*, 2009)

Recent studies suggest that transcriptional termination is linked to mRNA processing and may be required for optimal gene expression in human cells (West & Proudfoot, 2009) Consequently, transcription termination is now thought to be required for efficient mRNA processing, export, and therefore, for protein production.

Termination is important to pause RNA polymerase II (Pol II) and allow cleavage of the pre-mRNA with a functional polyadenylation [p(A)] signal at the p(A) addition site. However, the termination signal, located downstream of the p(A) signal, may also be required for efficient polyadenylation (Dye & Proudfoot, 2001). A failure in efficient termination by Pol II leads to inefficient pre-mRNA cleavage and polyadenylation, resulting in the formation of uncapped transcripts and consequently, low gene expression (West & Proudfoot, 2009).

Several models for polyadenylation associated transcription termination have been proposed:

- the allosteric model, or anti-terminator model, posits that p(A) signals cause a change in the complement of anti-termination factors associated with Pol II, making it more susceptible to termination (Logan *et al.*, 1987)

- the torpedo model suggests that a 5′ to 3′ exonuclease attacks the transcript at the p(A) cleavage point and degrades the nascent transcript in pursuit of the RNA polymerase, leading to termination (Connelly *et al.*, 1988)

- and more recently the combined allosteric/torpedo model (Luo W, et al., 2006), which includes a role for the phosphorylation state of the Pol II C-terminal domain (CTD) in termination and 3′ end formation (Ahn *et al.*, 2004).

6.2.3 *HBB* transcription terminator

The process of cleavage and polyadenylation of *HBB* transcripts has been shown to occur efficiently only following RNA self-cleaving activity located in the 3' flanking region 600-1200bp downstream of the p(A) site. These events are dependent on transcriptional termination by Pol II in this region (designated as "βterm"). Recent studies performed with plasmid vectors suggest that the inclusion of the βterm region downstream of the p(A) site increased mRNA and protein production 4-fold (West & Proudfoot, 2009). The authors report that the addition of the βterm region into a globin expression system increased the overall efficiency of pre-mRNA cleavage, polyadenylation and transcription termination.

Based on these studies we sought to augment the function of the *HBB* cassette within the GLOBE vector by incorporation of the βterm region at a position 3' of *HBB* (**Figure 6.1; GLOBE2)**.

6.3 LVs used in this study

6.3.1 GLOBE2

GLOBE2 was generated by Dr Sha Sha, a former member of the Gene Expression and Therapy Group, as follows; a 850bp βterm region genomic fragment with 850bp from the *HBB* polyadenylation site, was excised from a pBluescript plasmid vector with *Hpa*I and *Eco*RV restriction enzymes and inserted into the original GLOBE at the *Hpa* I site, upstream of the cppt element (**Figure 6.1**).

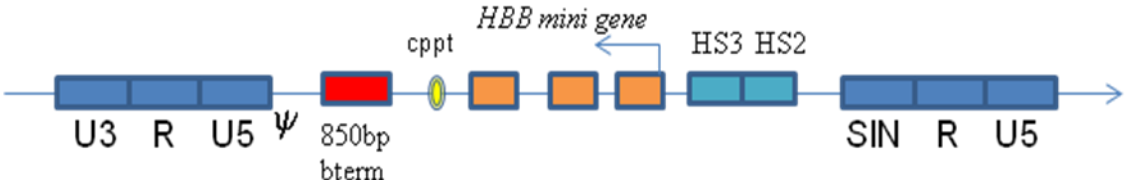


Figure 6.1 - Schematic illustration of GLOBE2 lentiviral vector. GLOBE2 is as the original GLOBE vector (Miccio *et al*., 2008) but with an 850bp genomic fragment spanning the *HBB* transcription termination region (βterm) inserted at a *Hpa*I site upstream of the LV cppt element.

Results obtained *in vitro* by transduction of murine erythroleukemia (MEL) cells and *in vivo* by *ex vivo* BM transplantation of HSC in mice using a eGFP reporter gene variant of GLOBE2, showed overall expression greater than 4-fold higher compared to a vector lacking βterm (Kao, V and Antoniou, M, *unpublished results*).

Based on these observations we decided to proceed with experiments *in utero* with the GLOBE2 LV.

6.3.2 Vector preparation

GLOBE2 vector was prepared as described in section **2.2.3**, Chapter 2. Viral stocks were serial diluted, used to transduce cultures of known cell density, and incubated for 3 days after which DNA was extracted. Extracted DNA was used to determine viral stock titre by qPCR analysis by comparison with a sample of known vector copy number.

Viral stocks of GLOBE2 ranged between 10^7-10^8 TU/mL. This titre value is relatively lower than that obtained with the equivalent eGFP-reporter construct and those routinely obtained with GLOBE. In all likelihood this is due to the larger size of GLOBE2, which closely approaches the 8kb size limit of LVs.

6.4 Animal experiments

6.4.1 *In utero* injections

Injections *in utero* were performed as previously described (section **2.2.4.2**, Chapter 2) and following the methodologies applied in experiments with eGFP and luciferase reporter vectors (sections **5.3.3** and **5.4.2**, Chapter 5; section **2.2.4.2**, Chapter 2). Based on the results obtained *in utero* with eGFP and luciferase reporter vectors, the whole set of injections in this study employing GLOBE2 were performed at embryonic day 14 (E14).

Unborn mice were injected with equal amounts of transgene vector and *Optimen* solution (control). After initially injecting 14 foetuses, 8 survived to birth and were studied post-natally.

6.4.2 Animal procedures

Blood was collected from live animals 120 days after in utero injection. As described in section **2.2.4.3**, Chapter 2. Mice were sacrificed by cardiac puncture so that the maximum amount of circulating blood was obtained, 250 days after *in utero* injection, as described in section **2.2.4.4**, Chapter 2. Blood samples were used for DNA extraction for proviral copy number determination (section **6.5.1.1**), haemoglobin-ELISA assay for quantification vector expression (section **6.5.1.2**) and the 250 day sample was used to estimate *HBB* expression levels in blood (section **6.5.1.3**)

Bone marrow cells were isolated from femurs and tibiae. Liver and spleen tissues were collected from carcases. Collected cells were used for proviral copy estimation (section **6.5.2.1**) and expression analysis (section **6.5.2.2**). DNA removed from tissues was used to estimate proviral copy number in mice tissues (section **6.5.3.1**).

6.5 Expression analysis

6.5.1 Analysis of peripheral blood

6.5.1.1 Determination of proviral copy number

DNA was extracted from peripheral blood and used for vector copy number estimation (section **2.2.5.1.1**, Chapter 2). Proviral estimation was determined using qPCR and vector presence and concentration in samples was estimated by comparison with an endogenous control and a sample of known copy number (section **2.2.5.2**, Chapter 2).

At the time of sacrifice, 2 mice not subjected to the injection procedure were added to the experimental set, namely mouse numbers 9 and 10. (Note: the control mice were injected with *Optimen* medium instead of viral vector stock).

Mouse number	Gender	LV copy number 120 days post injection	LV copy number 250 days post injection
1	male	**1.405**	**0.229**
2 control 1	male	**0.108***	**0.003****
3	male	**0.432**	**1.109**
4	male	**0.376**	**1.354**
5	male	**1.613**	**0.161**
6 control 2	male	**Undetected**	**Undetected**
7	female	**2.910**	**0.952**
8 control 3	female	**Undetected**	**Undetected**
9 un-injected mouse[1]	male	**-**	**Undetected**
10 un-injected mouse[1]	male	**-**	**Undetected**

Table 6.1. GLOBE2 LV copy number values in peripheral blood at 120 and 250 days post-*in utero* injection. Values were obtained by analysis of DNA by qPCR. The results show that GLOBE2 sequences are present in the mouse circulation. Copy number values averaged at 1.347 (120 days) and 0.761 (250 days) viral copies per cell. Mice injected with GLOBE2: 1, 3, 4, 5 and 7.

[1]Mice not subjected to the injection procedure. *The copy number estimation was repeated and this value was consistent (n=2)**.** The amplification cycle in this reaction was very high at 35-38 cycles (reaction with 40 cycles). Amplification at this stage of the reaction cannot be considered copy number amplification but residual traces of vector in the sample. ** The samples were run in duplicate and the PCR was repeated but a positive value (albeit at a very low level) was obtained for amplification of control mouse 1; analysis of the results obtained in the qPCR reaction show that this sample amplified DNA with the vector detection primers (labelled at LTR primers in section **2.2.3.4.2,** Chapter 2) at the 35th cycle, when the whole reaction was set to amplify 40 cycles.

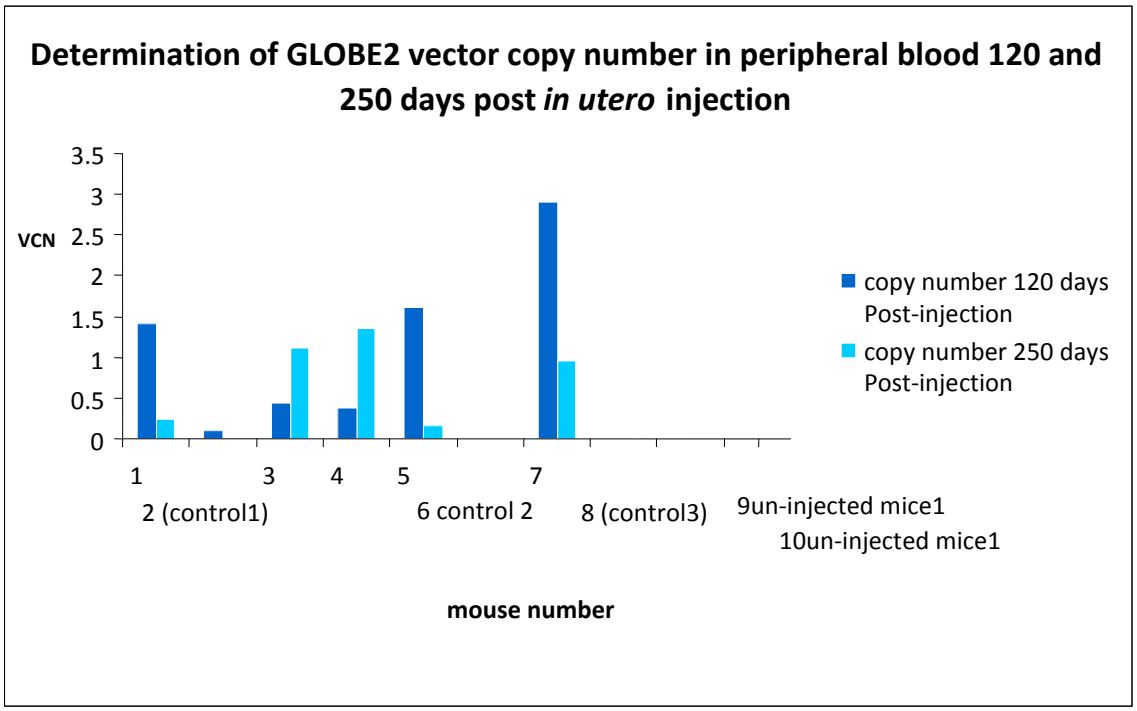


Figure 6.2 - Determination of GLOBE2 vector copy number in peripheral blood of mice at 120 and 250 days post-*in utero* injection (n=2). DNA extracted from peripheral blood collected from mice injected *in utero* with the GLOBE2 was analysed for vector copy number (VCN). Graphical representation of data shown in **Table 6.1**. Results clearly show vector presence in peripheral blood.

6.5.1.2 Determination of *HBB* chain levels in haemoglobin in peripheral blood

Determination of GLOBE2-derived *HBB* chain levels 250 days post-injection was performed by Haemoglobin ELISA. The technique takes advantage of structural differences between mouse and human, detecting the later (see section **2.2.5.5**, Chapter 2). *HBB* chain presence was estimated by comparison with a standard normal human sample and normalised with respect to vector copy number.

Mouse number	**Gender**	***HBB* chain levels 120 days post injection (ng/ml)**	***HBB* chain levels 250 days post injection (ng/ml)**
1	male	**0.153**	**0.0915**
2 control 1	male	**0.256**	**0.088**
3	male	**0.155**	**0.104**
4	male	**0.079**	**0.042**
5	male	**0.182**	**0.186**
6 control 2	male	**0.082**	**0.062**
7	female	**0.174**	**0.101**
8 control 3	female	**0.09**	**0.028**
9 un-injected [1]	male	**-**	**0.069**
10 un-injected [1]	male	**-**	**0.093**

Table 6.2 - GLOBE2-derived *HBB* chain levels in peripheral blood collected at 120 and 250 days after *in utero* injection (n=2). Blood samples were diluted and used to determine *HBB* chain levels in peripheral blood using the Human Hemoglobin ELISA Kit. [1]Mice not subjected to the injection procedure.

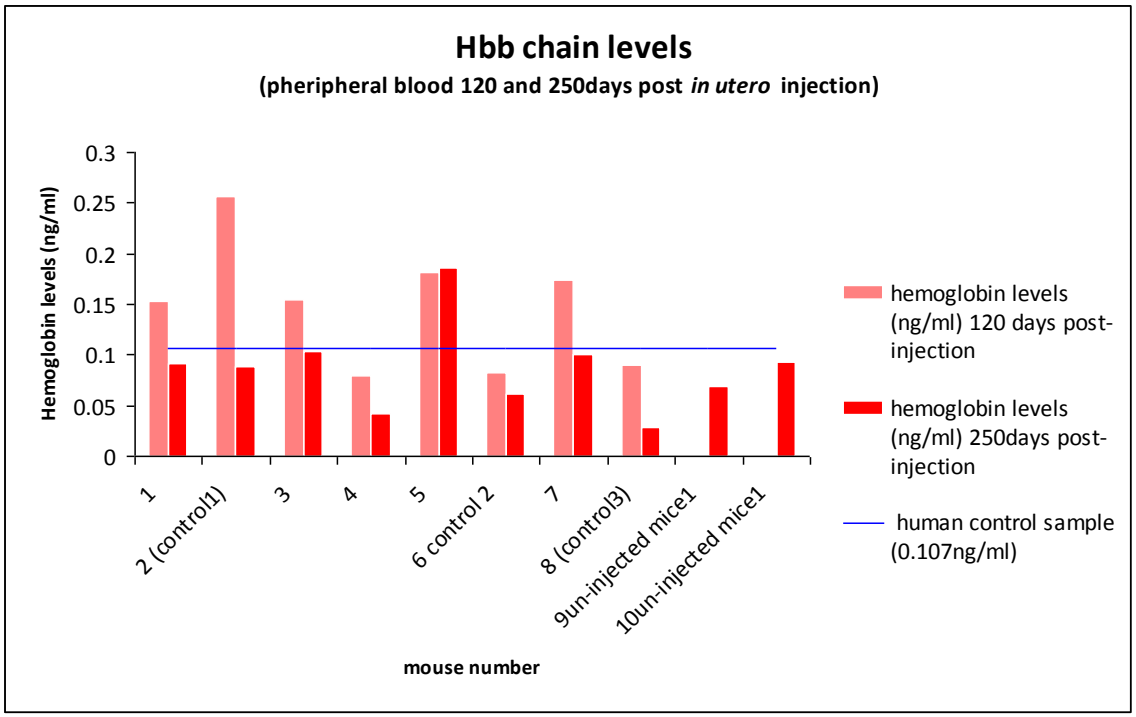


Figure 6.3 - *HBB* chain levels in peripheral blood of mice collected at 120 and 250 days after *in utero* injection. *HBB* chain levels obtained (ng/ml) for both time points were plotted against the haemoglobin value obtained with a human control sample (0.107ng/ml). Comparison of both assay values (120 and 250 days) by Student's t-test returned a P-value of 0.0517 (P-value≥0.05), indicating no differences between the 2 groups of results analysed. Samples were analysed in duplicate.

The value obtained with mouse number 2 (control 1) raised questions regarding the specificity of the assay. However, there is some consistency between the values of control 2 and 3 (0.082 and 0.09) and the negative (un-injected) samples also scored values below the human positive control level. In addition, the P-value obtained by the comparison of both set of results (120 and 250 days) in a Student's t-test shows consistency in the set of results.

We decided to include this set of results from a merely qualitative point of view, as the results show a trend and overall consistency, but this assay cannot be interpreted quantitatively.

The graph presented in **Figure 6.3** although informative hides individual variance and does not account for the amount of vector in peripheral blood (amount of vector that gives rise to *HBB chains* in circulation). Therefore, the results obtained from *HBB* chain analysis together with the values obtained for GLOBE2 proviral copy number were used to calculate *HBB* chain levels per copy number to obtain a more quantitative presentation of the results (**Figure 6.4**).

Analysis of the graph in **Figure 6.4** shows that *HBB* chain levels at 120 days post-injecting are constant and ranged from 0.2-0.4ng/ml per vector copy. At 250 days post-injection a greater variation of expression was observed with mouse number 5 showing the highest values of *HBB chains* per copy. Overall, the results show that over time *HBB* chain values in the peripheral circulation decreased in mouse numbers 3 and 4 but in mice 1, 5 and 7 surpassed the value obtained at 120 days post-injection.

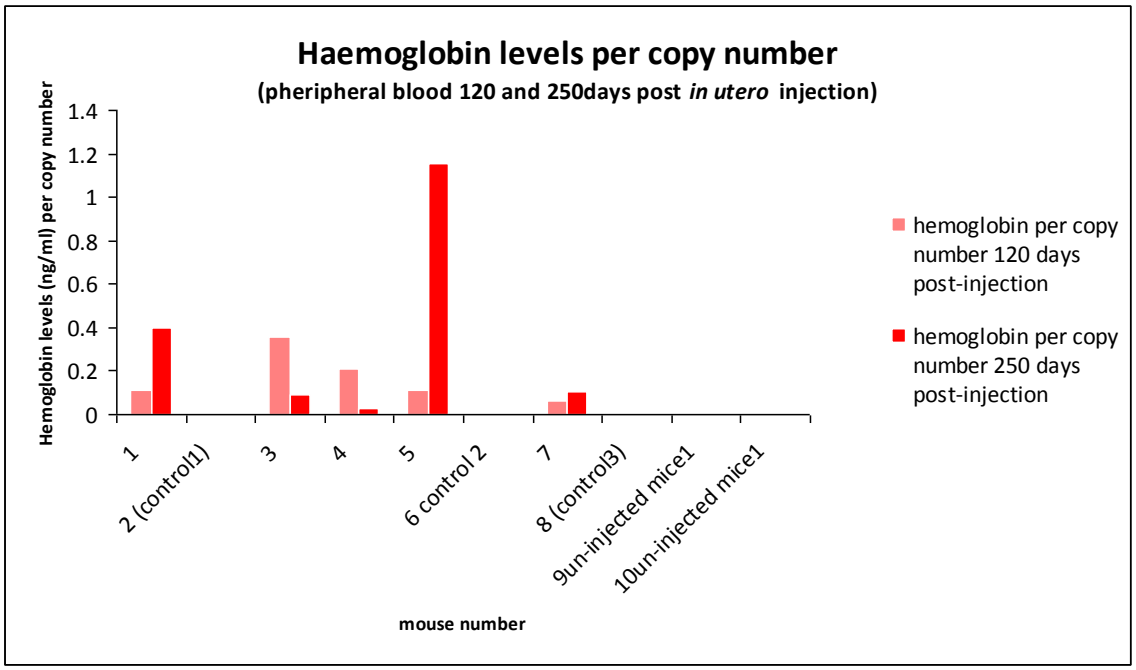


Figure 6.4 - *HBB* chain levels in peripheral blood per GLOBE2 copy number. *HBB* chain levels obtained from both time points (120 and 250 days post-in utero injection) (**Table 6.2, Figure 6.3**) were plotted against GLOBE2 copy number values (**Table 6.1, Figure 6.3**).

6.5.1.3 Quantification of *HBB* mRNA levels in peripheral blood

RNA extraction from peripheral blood was performed as described in section **2.2.5.1.2**, Chapter 2. For consistency, the same cell sample was used for simultaneous DNA and RNA extraction (respectively for copy number estimation and *HBB* mRNA quantification).

Figure 6.5 shows a qualitative analysis of the RNA samples by agarose gel electrophoresis. The result shows the presence of intact 28S and 18S ribosomal RNA bands in the correct 2:1 ratio confirming integrity of the sample preparations.

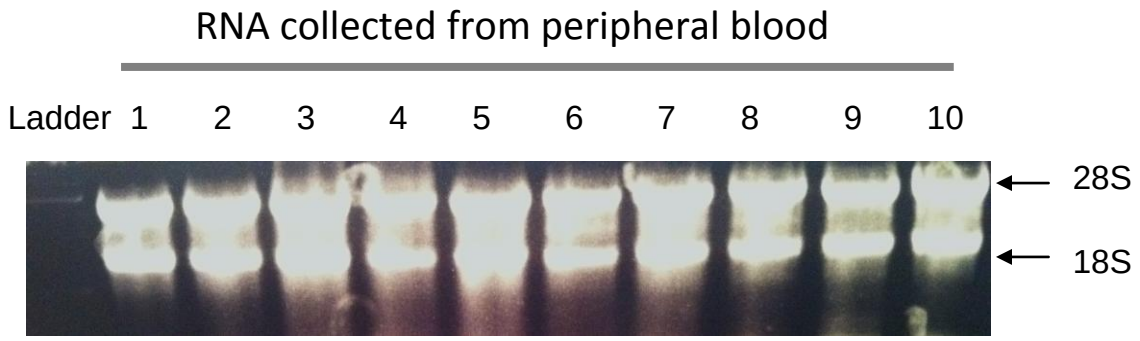


Figure 6.5 - Analysis of RNA extracted from peripheral blood, 250 days post-*in utero* injection by agarose gel electrophoresis.

6.5.1.3.1 Reverse Transcriptase (RT) qPCR

Total RNA was converted into complementary DNA (cDNA) by means of a reverse transcriptase reaction using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) kit (*Invitrogen*; see section **2.2.5.3.2**, Chapter 2).

Quantification of *HBB* expression levels from GLOBE2 per mouse was achieved through SYBR® Green qPCR. **Figure 6.6** shows *HBB* gene expression per vector copy in relation to endogenously expressed *GAPDH*.

The results obtained were analysed using the $2^{\Delta\Delta Ct}$ methodology (Winer *et al.*, 1999, *Termo Scientific Tech notes*, Josh Haimes & Melissa Kelley, 2010; Applied Biosystems *Guide to perform relative Quantitation of Gene Expression Using Real-Time Quantitative PCR*, 2008) using untransduced controls as calibrators to determine *HBB* level of expression per copy number.

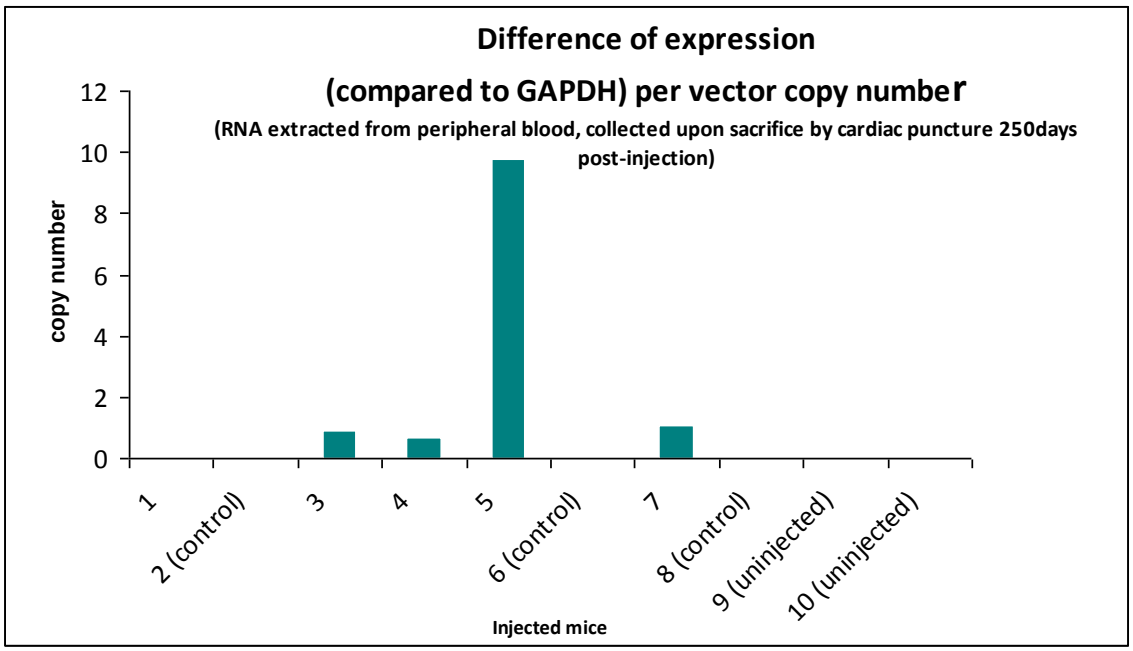


Figure 6.6 – Difference of \HBB expression *HBB* transgene expression compared to GAPDH per vector copy number in peripheral blood (n=2). Total RNA extracted from blood collected 250 days post-*in utero* injection was analysed to determine *HBB* transgene expression. RNA samples were equalized with regards to their concentration using a Nano-Drop 1000D detection system and reverse transcribed into cDNA using the M-MLV RT kit. cDNA sample quantification was determined by qPCR and plotted against peripheral blood copy number values (*section* 6.6.3.1) to determine expression per vector copy. Mean fold difference values were obtained by dividing by half taking into account GAPDH endogenous gene copy number of 2 per cell versus integration of the transgene at 1vector per cell.

6.5.2 Bone Marrow

6.5.2.1 Determination of proviral copy number from DNA extracted from bone marrow cells

Mice number	gender	Estimated vector copy number (BM, 250 days post injection)
1	male	**0.114**
2 control 1	male	**Undetected**
3	male	**0.892**
4	male	**0.359**
5	male	**0.036**
6 control 2	male	**Undetected**
7	female	**0.454**
8 control 3	female	**Undetected**
9 un-injected mouse[1]	male	**Undetected**
10 un-injected mouse[1]	male	**Undetected**

Table 6.3 - GLOBE2 vector copy number in BM cells analyzed 250 days after *in utero* injection (n=2). Vector copy number determined by qPCR ranged between 0.03-0.89 (average 0.371) per bone marrow (BM) cell. [1]Mice not subjected to the injection procedure.

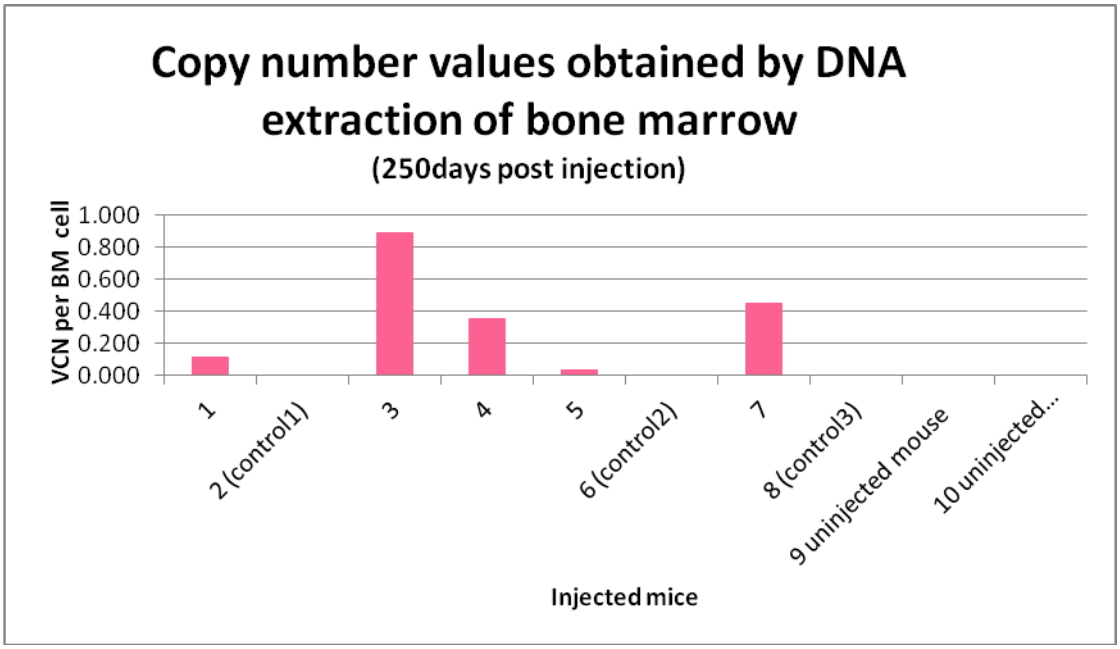


Figure 6.7 - Determination of GLOBE2 vector copy number in bone marrow cells 250 days post-*in utero* injection (n=2). DNA extracted from bone marrow cells from mice injected *in utero* with the GLOBE2 was analysed for vector copy number (VCN) per cell by qPCR.

With the exception of mouse number 3 (0.892), all mice injected with vector showed copy number values in BM lower than 0.5 copies per cell, but still, all at higher values compared with the control mice (numbers 2, 6 and 8).

6.5.2.2 Quantification of *HBB* mRNA levels in bone marrow cells

The same cell sample was used for simultaneous DNA and RNA extraction (respectively for vector copy number estimation and *HBB* mRNA quantification). **Figure 6.8** confirms the integrity of the RNA extracted. The samples were resolved by agarose gel electrophoresis and show the presence of the 28S and 18S ribosomal RNA bands in the desired 2:1 ration.

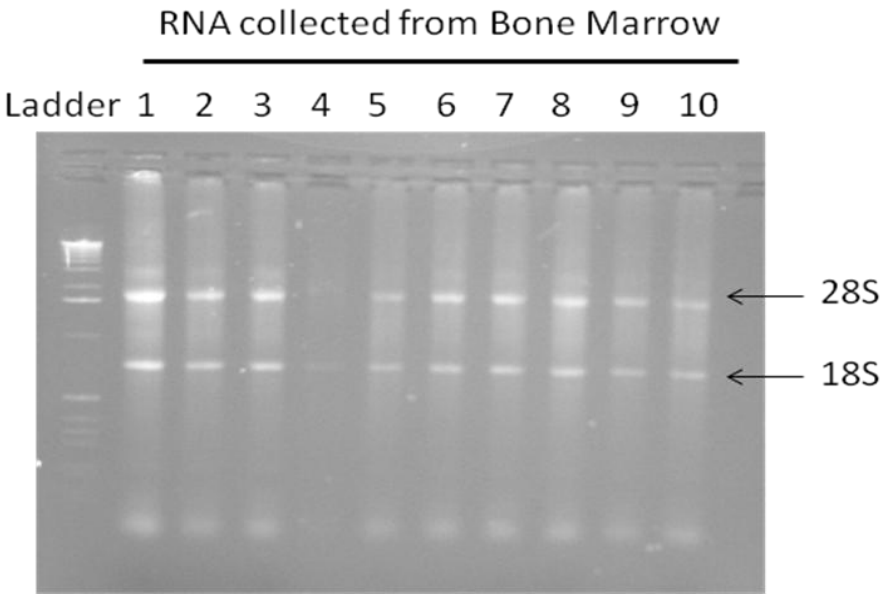


Figure 6.8 - RNA extracted from bone marrow cells at 250 days post *in utero* injection. The picture shows analysis of the RNA samples by agarose gel electrophoresis, which shows intact 28S and 18S ribosomal RNA bands.

Spectrophotometric analysis indicated a high level of purity ($_{A260/280nm}$≥1.7). Total RNA was converted into complementary DNA (cDNA) by means of a reverse transcriptase reaction using the Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) kit (*Invitrogen*; see section **2.2.5.3.2**, Chapter 2).

Quantification of GLOBE2 *HBB* expression levels per mouse was achieved through SYBR® Green qPCR. **Figures 6.9** show *HBB* expression per GLOBE2 vector copy in relation to endogenously expressed *GAPDH*.

Results obtained were analysed using the $2^{\Delta\Delta Ct}$ methodology (Winer *et al.,* 1999) and *Termo Scientific Tech notes* (Josh Haimes & Melissa Kelley, 2010), using untransduced controls as calibrators to determine *HBB* expression level per vector copy number.

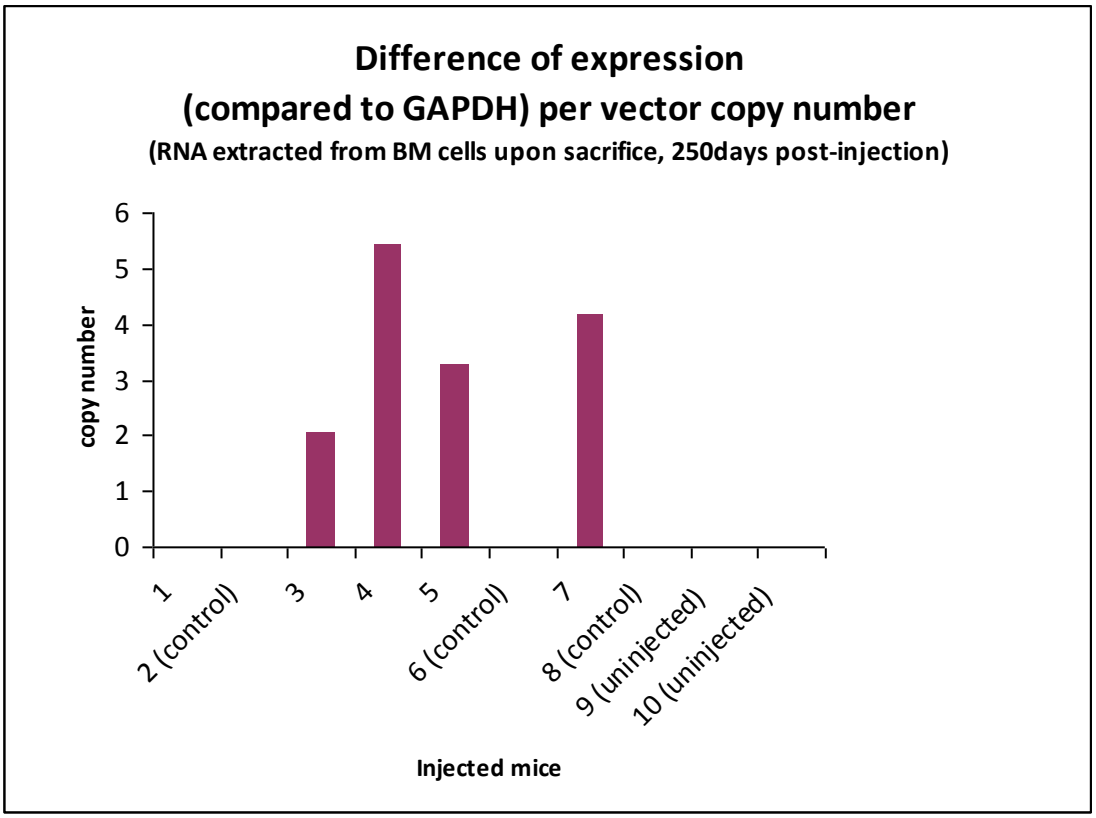


Figure 6.9 - *HBB* mRNA levels per GLOBE2 vector copy number in bone marrow (n=2). RNA was isolated from bone marrow (BM) cells at 250 days post-*in utero* injection and analysed to determine *HBB* transgene expression. RNA samples were equalised with regards to their concentration using a *NanoDrop* 1000D detection system and reverse transcribed into cDNA. *HBB* mRNA quantification was determined by a qPCR and plotted against vector copy number (see *section* 6.6.3.2) to determine mRNA expression per vector copy as a fold difference over GAPDH values.

6.5.3 Analysis of liver and spleen tissues

6.5.3.1 Determination of proviral copy number from DNA extracted from liver and spleen at 250days post-injection

Mice number	gender	Vector copy number Liver (250 days post injection)	Vector copy number Spleen (250 days post injection)
1	male	**2.457**	**0.152**
2 control 1	male	**Undetectable**	**Undetectable**
3	male	**3.142**	**0.174**
4	male	**2.561**	**0.25**
5	male	**1.925**	**0.339**
6 control 2	male	**Undetectable**	**Undetectable**
7	female	**2.469**	**0.21**
8 control 3	female	**Undetectable**	**Undetectable**
9 un-injected mouse[1]	male	**Undetectable**	**Undetectable**
10 un-injected mouse[1]	male	**Undetectable**	**Undetectable**

Table 6.4 - GLOBE2 vector copy number in liver and spleen tissues at 250 days post injection. Total genomic DNA from liver and spleen was analysed for GLOBE2 LV copy number by qPCR. Liver average copy number is 2.51; spleen average copy number is 0.225. [1]Mice not subjected to the injection procedure.

6.6 Conclusions

6.6.1 Analysis of GLOBE2 vector copy number in peripheral blood

Analysis of GLOBE2 vector copy number values obtained by DNA analysis and comparison with a control cell sample with a known single copy integration event, revealed that mouse numbers 1 and 5 showed values close to 1 copy per cell, whilst mouse number 7 showed values close to 3 integration events per cell. Mouse numbers 3 and 4 contained less that 1 copy per 2 cells in peripheral blood with values of 0.4327 and 0.3762 respectively. The initial vector copy number values at 120 days post-injection (**Table 6.1, Figure 6.2)** are encouraging and confirmed GLOBE2 vector presence in the circulation. However, at this point it was still unclear whether this vector resulted from progenitor cells in the circulation upon injection or newly derived blood constituents from HSCs in BM.

The analysis of vector copy number values in peripheral blood at 250 days post-*in utero* injection (**Table 6.1, Figure 6.2**) show that mouse numbers 3 and 4 have the highest amount of vector in the bloodstream with 1.909 and 1.354 vector copies per cell respectively. Mouse number 7 showed a vector copy number of 0.952/cell whereas mouse numbers 1 and 5 showed the lowest level of vector marking at 0.229 and 0.161 per cell respectively.

The comparison of the 120 day and 250 day vector copy number estimations shows that mice with higher copy number values at 120 days post-injection decreased in their vector copy number in the circulation over time. In contrast, mouse numbers 5 and 6, which showed relatively low viral copies per cell in the peripheral blood at 120days, showed a significant increase in vector copies by 250 days.

Based on these results it is possible that the vector copy values obtained initially for peripheral copy number at 120 days post-injection may be arising from transduced progenitor cells rather than HSC.

6.6.2 *HBB* chain levels in blood

Determination of GLOBE2 vector derived HBB chain levels at 120 days post-injection was performed by Haemoglobin ELISA (**Table 6.2, Figure 6.3**).

Upon careful consideration, we decided to analyse the results obtained in this set of experiments excluding mouse number 2, as the value obtained for this sample differs significantly from the values obtained with the other mice analysed, the other controls (mice 6 and 8) and the values obtained for the negative controls. With this in mind, the values obtained correlate with the values obtained for vector copy number; mouse numbers 1, 5 and 7 showed highest values of HBB chain levels, as can be predicted from the values obtained for vector copy number (1.4055, 1.6138 and 2.9104 respectively). Mouse number 4 showed a HBB chain value very similar to control mice (6 and 8), as can be expected by the value obtained for proviral copy number estimation in peripheral blood (0.3762).

The surprising result in this assay is for mouse number 3, which shows steady HBB chain levels (0.155ng/ml) above the human control sample value (0.107ng/ml) but with a proviral copy number in blood of only 0.4327 viral copies per cell. This implies that accumulation of HBB chains in the circulation increases whilst vector copy number is gradually decreasing. The discrepancy between vector copy number and HBB chain levels in this mouse may be accounted for by preferential expansion of a clone or clones of HSC or myeloid precursor cells with high level expressing integration events.

The specificity of the human HBB chain assay was once again called into question upon analysis of the blood samples collected at 250 days post-injection as control mice injected with solely *Optimen* medium (mice 2, 6 and 8) scored surprisingly high values (**Figure 6.3)**.

We nevertheless decided to include this set of results in this study because, even though the control samples showed positive values, these are generally lower than the values obtained with samples from animals injected with GLOBE2. In addition, when analysed purely qualitatively these data provide evidence for the presence of HBB chains in the circulation of the vector injected animals. In an attempt to clarify this issue upon sacrifice and second collection of samples we also isolated RNA samples for the determination transgene expression levels though qPCR methodology.

In order to account for individual differences in the samples due to variation in GLOBE2 vector copy number, we plotted the values obtained for *HBB chains* against the values found for proviral copy number in circulation (**Figure 6.4**). (Note: in this figure we omitted the positive values obtained for mouse 2).

Overall, the results show that over time HBB chain values in the peripheral circulation decreased in mouse numbers 3 and 4 but in mice 1, 5 and 7 surpassed the value obtained at 120 days post-injection. The variation between the results obtained at both time points can be accounted for by the degree of GLOBE2 vector presence in the BM of the injected animals; that is, BM engraftment of transduced HSCs will cause differentiated erythroid cells to produce HBB chains, whereas transduction of only committed progenitor cells, the transgene will not be present in BM HSCs and vector presence in the circulation will diminish over time.

6.6.3 HBB mRNA levels in peripheral blood at 250 days post-injection

With the exception of mouse 5 (see below), injected mice showed consistent HBB transgene mRNA values in peripheral blood (approximately 1-fold higher than the GAPDH control) whilst control mice 2, 6, 8, 9 and 10 did not produce a detectable qPCR product (**Figure 6.6)**.

6.6.4 Peripheral blood - comparison of results obtained

There is generally a good correlation between transgene mRNA levels per vector copy number (**Figure 6.6**) with the values obtained for HBB chain presence in peripheral blood cells determined by haemoglobin ELISA. Overall both studies show complementary results with the transgene mRNA analysis validating the HBB chain levels in the blood cells analysed. For mouse number 5, both experimental sets of results showed high expression for the vector when corrected for copy number values obtained in peripheral blood.

The surprising finding in this comparison is mouse number 1, which showed a positive level of HBB chains per vector copy number but no transgene mRNA expression.

Based on results obtained previously with LVs containing *HBB* expression cassettes, including within human clinical trials (Cavazzana-Calvo *et al.,* 2010), it would be interesting to determine the sites of vector integration in mouse 5 to determine if the high expression seen in this case is due to clonal expansion of a specific cell population caused by host genome elements near the site of insertion causing upregulation. However, a study with the original GLOBE vector clearly showed that this integrates preferably within transcriptionally active genes, with no bias towards integration in proximity of oncogenes, as reported to occur with both lentiviral and retroviral vectors (Roselli *et al.,* 2010). It is thus important to notice that this mouse on sacrifice showed no evidence of tumours or other illness.

6.6.5 Analysis of GLOBE2 vector copy number in bone marrow

The existence of positive engraftment of cells harbouring GLOBE2 in BM validates the initial hypothesis formulated in this study and shows that it is possible to transduce HSCs whilst present in the foetal liver at a pre-migration stage into BM niches.

Results obtained by analysis of DNA extracted from BM cells (**Figure 6.7**) show that, although at a relatively low copy number and with less than 1 copy per cell in the BM compartment, GLOBE2 vector is present.

At 250 days post-injection it is possible to see a clear correlation between BM engraftment (copy number values obtained in BM cells) and the viral copies present in the mouse bloodstream (copy number in peripheral blood, **Figure 6.2**). These results clearly suggest that the vector copy number found in the circulation is highly dependent on progenitors arising from transduced cells in the BM.

6.6.6 *HBB* mRNA expression levels in BM at 250 days post-injection

Results obtained from quantification of *HBB* transgene expression in BM cells (**Figure 6.9**) show that with the exception for mouse number 1 all other injected mice gave a positive result compared to the GAPDH normalised gene.

Given that the values obtained for copy number estimation in peripheral blood were much higher than the values obtained in BM cells at 250 days post-injection (**Figures 6.2** and **6.7**), it is possible to conclude that transduced cells in BM are expressing higher levels of transgene than those in the circulation. This may be accounted for by the presence of larger numbers of HBB mRNA positive erythroid progenitor cells in BM compared to peripheral blood.

6.6.7 Analysis of copy number in non-haematopoietic mouse tissues at 250 days post-injection

As expected the liver and spleen in GLOBE2 injected mice showed vector presence. Vector copy number is higher in the liver due to LV delivery *in utero* is performed via the vitelline line yolk sac vessel, which drains into this organ.

6.6.8 Overall analysis

In this study, we investigated the possibility of an intrauterine gene therapy approach for correction of the haemoglobinopathies in wild type mice. In an overall analysis of the results, it is possible to conclude that *in utero* injection with GLOBE2 (**Figure 6.1**) is capable of transducing the HSC population in the foetal liver resulting in subsequent engraftment in BM until, at least, 250 days post-injection. This engraftment is responsible for the observed prolonged transgene presence and expression at both an mRNA and protein level in peripheral blood.

6.6.9 Conclusion remarks

- The results show positive engraftment of BM cells containing the GLOBE2 vector following *in utero* injection

- The results validate the initial hypothesis formulated in this study and show that it is possible to transduce HSC while in the foetal liver, before migration into the BM niches

- Vector presence in peripheral blood and BM up to at least 250 days after injection suggests targeting of HSCs in foetal liver

- *HBB* transgene expression levels determined by mRNA and HBB chain analysis suggest that levels capable of rescue of the β-thalassaemia phenotype are possible and pave the way for a therapeutic experiment in mouse models of β-thalassaemia using the GLOBE2 vector.

Chapter Seven

DISCUSSION

The last decade has seen great advances in the field of gene therapy with the first successful outcomes in clinical trials (Cartier *et al.,* 2009; The use of gammaretroviral vectors has provided long-term correction for severe combined immune deficiency (SCID) following *ex vivo* delivery to HSC in conditions such as SCID-X1 (Cavazzana-Calvo *et al*., 2000; Gaspar *et al*., 2004) and SCID-ADA (Aiuti *et al.,* 2009). AAV vectors have been employed with success in targeting the retinal pigment epithelium in Leber's Congenital Amaurosis (Bainbridge *et al.,* 2008; Maguire *et al.,* 2008; Ali *et al.,* 2012). More recently outcomes from the use of lentiviral vectors in clinical trials to treat inherited monogenic diseases have been reported. Using *ex vivo* delivery to HSC a successful outcome was achieved in patients with X-linked adrenoleukodystrophy (Cartier et al., 2009) and a partial success in one case of β-thalassaemia (Cavazzana-Calvo *et al*., 2010).

Despite these successes, problems regarding insertional mutagenesis resulting from viral vector integration remain and the search for an optimal delivery vector is far from its end.

In 2003 the follow up from the X-linked SCID human trial in Paris reported that one of the patients had developed a monoclonal lymphoproliferative disease and since then, four additional patients have developed similar forms of leukaemia (Hacein-Bey-Abina *et al*., 2003). Analysis of T-cell clones from patients revealed integration near the LMO2 proto-oncogene in four out of the five cases, leading to aberrantly high transcription (Hacein-Bey-Abina *et al*., 2003a).

More recently, there has been reports of the apparent successful treatment of an 18-year old patient with transfusion dependent HbE/β^0-thalassaemia via BM re-implantation of genetically modified HSC with an LV containing a functional *HBB* (Cavazzana-Calvo *el al*., 2010). Results showed 10-20% of reconstituted HSCs but it was not until further analysis confirmed that 50% of genetically modified cells and consequent increased levels of *HBB* seen in this patient, were a result of clonal dominant expansion of a single myeloid progenitor harbouring an LV integration within *HMGA2*, a

proto-oncogene. This integration event resulted in an exons 1-3 truncated HMGA2 peptide being produced and abolition of the miRNA-mediated control of this gene normally present in the terminal exon 5. Once again a therapeutic success is associated with a (fortuitous) insertional mutagenesis event.

As things stand, despite key milestones in the field of gene therapy being achieved, there is a wide window for improvement in LV design still to be developed, mainly with regards to safe integration, increased transduction efficiency and transgene expression levels. This project addresses the major issues associated with gene therapy lentiviral cassettes targeting the haemoglobinopathies and aims to develop a therapeutic globin vector capable of overcoming silencing and to achieve sustained, long term, tissue specific and reproducible expression with minimum genome disturbance.

7.1 Comparative study accessing the ability of both ubiquitous and tissue specific elements to express from within an ID-LV context

In order to address problems associated with integration mutagenesis ID-LV vectors have been developed (Esposito and Craigie, 1998; Gerton *et al*., 1998; Leavitt *et al*., 1996; Masuda *et al*., 1998). Whilst retaining attractive properties of LVs, such as the ability to transduce both dividing and quiescent cells, ID-LV are delivered with a mutant integrase protein, which specifically prevents proviral integration resulting in the generation of increased levels of 1-LTR and 2-LTR circular episomes in transduced cells.

When compared to regular integrating LVs, ID-LVs have significantly lower risk of causing insertional mutagenesis together with lower risk of generating replication competent retroviruses. So far, success from within an ID-LV context has solely been reported with transgene expression from ubiquitous promoter/gene combinations (Yanez-Munoz *et al*., 2006; Apolonia *et al*., 2009*)* using vectors with a mutation in the catalytic motif of the integrase protein (DD35E, residue D64) that impairs integration (Leavitt *et al*., 1996).

The first question in this study (Chapter 3) was whether expression from within ID-LV can be achieved only with ubiquitous regulatory elements or whether this can be extended to tissue specific control.

Tissue-specific promoters are usually relatively weak when compared with ubiquitous promoters but their function is almost invariably augmented by enhancer or LCR elements to boost expression in a cell-specific manner. Vectors containing such tissue specific regulatory elements augmenting transcription offer many significant advantages over non-specific vectors and are able to limit immune responses against the therapeutic product due to abrogation of ectopic expression in antigen presenting cells (Weeratna *et al.*, 2001). Furthermore, they have less potential to promote insertional

oncogenesis and most importantly limit ectopic expression that may have toxic consequences for other organs.

Tissue-specific elements from the human *HBB* LCR have been successfully included in the development of LVs for the treatment of the haemoglobinopathies (see section 1.4.1) with LCR sites HS2 and HS3 included in our therapeutic GLOBE vector (**Figure 1.18**).

In Chapter 3 we document the comparison of both ubiquitous expressing and tissue specific elements in an ID-LV using the same integrase mutation as previously described (D64; Yanez-Munoz *et al.*, 2006). All vectors expressing from either ubiquitous (SFFV) or tissue specific (*HBB* and muscle *DES* / *CK-M* specific) regulatory elements were produced without major complication and expressed from both IP an ID configurations. Thus, we demonstrate that utility of ID-LV can be extended to provide tissue-specific as well as varying degrees of ubiquitous therapeutic gene expression depending on the control elements used.

As initially expected, the expression for ID-LV constructs decreased significantly following successive rounds of cell division and reached values close to zero in some vector configurations. Nevertheless, this study demonstrates that tissue specific as well as ubiquitous promoters can express from an ID-LV context, albeit at low intensity levels.

Our results also demonstrate that the concentration of initial viral DNA (3 days post transduction) is similar for all integrating and non-integrating vectors analysed, so it is possible to conclude that the studied integrase defective vectors do not have an impact on reverse transcription. However, the sustainability of the vector in mitotic cell populations is dependent on initial MOI. Because transduction at high copy number can be impracticable as large scale production of virus is expensive and laborious, the approach is limited to transient expression. Also, the transduction of large amounts of a given vector aiming for initial high copy number values may lead to over expression of

the transgene, protein accumulation and toxicity. Nevertheless, since maintenance and sustained expression from ID-LV can occur in non-dividing cell populations (Yanez-Munoz *et al.,* 2006), it is possible that long-term tissue specific expression can also be obtained with this system in post-mitotic tissues.

7.1.1 Applications of ID-LVs

As in previous studies we found that ID-LVs are capable to deliver viral DNA to target cells, but regardless of the tropism of the driven vectors, expression from these vectors is transient due to vector loss through successive rounds of mitosis. Although limiting, there are several applications of such vectors. An example would be the development of cascade effects for cellular protein expression. For example, the expression of HoxB4 has been shown to increase self-renewal and promote expansion of HSCs (Antonchuk *et al*., 2002). However, permanent production of HoxB4 protein is associated with toxicity (Zhang *et al*., 2007b; Krosl *et al*., 1998). The use of ID-LVs in such a project would be of great use to provide temporary expression of the protein, avoiding effects of protein accumulation caused by LV integration.

A development of replicating ID-LV would be of great use in the field. This would simultaneously reduce the risk of insertional mutagenesis and maintain transgene expression in both, dividing and non-dividing cells. Previous studies have shown that it is possible to incorporate the simian virus 40 (SV40) OriT into ID-LVs allowing the maintenance of the episomal DNA in dividing cells, without reversion to an integrative phenotype. Replication and expression from this vector was sustained for 56 days in cells expressing the SV40 Large T-antigen. However, under the selective conditions of T-antigen, stable expression was not seen (Lu *et al*., 2004; Vargas, Jr. *et al*., 2008). Nonetheless, it is important to notice that these vectors still carry a risk of mutagenesis due to the presence of the SV40 Large T-antigen. An alternative system that can potentially confer replication status on ID-LV is the incorporation of a scaffold/matrix

attachment region (S/MAR) as has been achieved in plasmid vectors (Hagedorn *et al.*, 2011). However, to the best of our knowledge, success incorporating an S/MAR in an ID-LV has not been reported.

More recently, ID-LVs have also been shown to be efficient immunogens with long-lasting immune responses (Karwacz *et al*., 2009; Negri *et al*., 2007). In fact, these vectors make use of the infectivity properties of regular LV vectors efficiently stimulating cell-mediated and humoral immunity.

Finally, some groups have recently reported the use of ID-LVs to deliver specific DNA sequences without genome integration, but using a different strategy for stable expression in dividing cells. The first obvious development that comes to mind is the use of Rep protein from AAV. Rep protein can site-specifically integrate DNA substrates into the human genome, including circular DNA (Kotin *et al*., 1990). Therefore, it would be interesting to study the combination of this protein with the delivery of the target DNA mediated by ID-LVs.

DNA delivered by ID-LVs has been used as a template for homologous recombination. Examples include the phage ΦC31 integrase (Calos *et al.*, 2005), and the Sleeping Beauty transposase (Staunstrup *et al.*, 2009; Vink *et al.*, 2009).

To the best of our knowledge, the work presented here is the first of its kind to compare a range of tissue-specific and ubiquitous expressing vectors in an ID-LV context.

7.2 Integration proficient LV, the challenge

The study of ID-LVs, although informative, failed to provide a positive direction for this study. We therefore decided to focus this project in the improvement of current integrating LVs for the haemoglobinopathies.

Current gene therapy approaches in the search of an optimal vector as a therapeutic agent for the haemoglobinopathies highlights the need for a design of safer vectors along with a reduction in vector copy number in target cells, ideally limited to 1-2 copies per cell. However, the need of delivery of fewer vector copies brings the need for high-expressing vectors. Sustained therapeutic levels of protein expression with a low vector copy number can only be attained by the optimisation of transgene cassettes that display a low probability of silencing at most integration sites and efficient post-transcriptional processing.

The *HBB*-based GLOBE vector has shown successful transduction of HSC and correction of the thalassaemia phenotype in mice (Miccio *et al.,* 2008). Furthermore, GLOBE restores globin chain imbalance following transduction of CD34+ cells from thalassaemic patients (Roselli *et al.,* 2010). Analysis of integration sites revealed that GLOBE integrates preferentially, as expected, within transcriptionally active genes, with no bias towards integration in proximity of oncogenes. More importantly, despite integration within or near active genes, GLOBE did not give rise to increased host gene expression suggesting a low insertional mutagenesis potential (Roselli *et al.,* 2010). Modifications of of GLOBE made by the host lab have led to improvements with regards to vector expression levels mediated by inclusion of the βterm element (GLOBE2, **Figure 6.1**). This design is based on studies conducted with an HBB-eGFP-βterm cassette in LVs, which *in vivo* showed overall expression over 4-fold higher compared to to the same vector without βterm (Kao V *el al., unpublished results*).

7.2.1 Vector test studies: injection in neonatal mice with different reporter gene cassettes.

Neonatal injections offer a good model system within which to evaluate therapeutic gene therapy. We were inclined to use a UCOE-driven construct since this system has been proven to provide reproducible and sustained expression *in vivo* following *ex vivo* manipulation of HSC in mice (Zhang *et al*., 2007 & 2010). Nevertheless, we decided to conduct a comparison of different vector designs to determine which would be most suitable to take forward with *in utero* delivery experiments.

In Chapter 4 we document the expression obtained from the eGFP reporter gene in tissues of neonate mice injected via the temporal vein. As expected, expression at a high level was found in the liver (**Figure 4.6**) and spleen with none being detected in either lung or heart tissues. Expression in the spleen shows that vector is present in the peripheral blood system.

In this initial study we considered a couple of promoter /enhancer combinations to be tested:

Firstly, the CMV promoter is frequently used in an LV context (Follenzi *et al.*, 2000) and it is known by its strong promoter capability. In this set of experiments we were not able to monitor expression initially following injections but until sacrifice. However, it would be interesting to compare the initial expression of the CMV vector and then at periodic intervals post delivery. CMV is known to function as a strong promoter in certain tissues, stronger than the A2UCOE. However, the results obtained here showed CMV to express at lower levels than the A2UCOE, even though initial vector titre at the time of delivery was similar.

Secondly, the A2UCOE element was chosen to be incorporated in this study as it is not prone to transcriptional silencing (in contrast with strong viral promoters such as

SFFV and especially CMV) (Zhang *et al.,* 2007). UCOEs are able to fully negate the process of PEV and confer stable, high level transcription on a transgene in both stably transfected cell lines (Antoniou *et al.,* 2003; Williams *et al.,* 2005.) and from lentiviral-transduced cells (Zhang *et al.,* 2007 & 2010; Brendel *et al.,* 2011). . In addition, the A2UCOE has been shown to lack classical enhancer function, thereby reducing the risk of activating a nearby promoter, which is highly significant for improving the safety profile of integrating vectors.

We expected that a vector containing a WPRE sequence would, in theory, provide higher vector production and levels of expression per vector copy, but that comparison with and without this element needed to be performed by conducting *in vivo* studies in our neonatal model system. Results with both A2UCOE driven vectors +/-WPRE were consistent in the sense that they provided stable expression of the construct. It is visible by the pictures shown in **Figure 5.10** that both the A2UCOE-eGFP and A2UCOE-EGFP-WPRE show expression in most cells in the sample. Cells expressing A2UCOE-eGFP are less bright but expression from this vector is consistent in all cells in the sample shown. This is an important measure with regards to therapeutics in gene therapy. The ideal gene therapy vector transduces efficiently and expresses in all transfected cells, avoids variegation of expression, but is a concise and reproducible vector.

According to our data the vector that best suited the requirements of an 'ideal *in utero* vector' was the UCOE-eGFP-WPRE and was thus the vector used for further *in utero* marking experiments (Chapter 5).

These experiments prove the feasibility of the neonatal technique to deliver LV into the mouse circulation. Neonatal tests checked the feasibility of the vector design for the study and paved the way for the actual *in utero* experiments.

7.2.2 Novel therapeutic options: *in utero* gene therapy

The second project objective aimed to investigate whether a sustained and reproducible level of expression can also be obtained when an LV expressing human *HBB* is delivered *in utero* into wild-type mice and ultimately, if sufficient expression is obtained to attempt rescue of a knock-out model of β-thalassaemia by this approach.

In the last two Chapters in this study, we covered *in utero* delivery (Chapters 5) and accessed whether this approach is feasible to be applied as a therapeutic approach for the haemoglobinopathies (Chapter 6).

The *in utero* procedure for therapeutic gene delivery has been developed by others previously (Dejneka *et al.,* 2004: Rucker *et al.,* 2004: Waddington *et al.,* 2004) and therapeutic progress in animal models using this technique has been reported in a broad spectrum of genetic disorders. The approach provides the correction of the disease phenotype in time to prevent manifestations of disease and, most importantly, the vector delivery is thought to occur during a period of low immunogenicity and therefore the technique prevents a severe immunological reaction that may to occur if the transgene is delivered later in the patient's life in some cases. However, the technique is still relatively new as its possible complications and limiting factors must be accessed before being used for clinical trials.

In utero studies performed so far have focused on liver transduction (Waddington *et al.,* 2004) and vector systemic delivery (van der Wegen *el al.,* 2006) and to our knowledge, the technique has not been used in marking studies attempting to target HSCs. Unlike effective fetal treatment of metabolic diseases, where liver cells could be targeted during most of the duration of pregnancy, treatment of hematopoietic disorders may have a much narrower window for effective HSC targeting. With this in mind, our study started with marking experiments to evaluate the ideal timing of *in utero* delivery

in mice that would account for the higher percentage of engrafted HSCs. We embarked on an experiment to compare different *in utero* delivery time points and check the protein/vector copy number value obtained in various organ compartments (Chapter 5).

As described previously, the A2UCOE-GFP-WPRE vector was used as an eGFP reporter gene marker in this phase of the *in utero* study. Injections *in utero* with this vector proved the feasibility of the experiment and showed positive results (Chapter 5). The vector showed sustained and reproducible expression of eGFP injected *in utero* into wild-type mice at both time points (E14, E16) analysed. However, staining with different antibodies both at E14 and E16 time points showed different marking profiles, with the E14 seeing a higher population of myeloid cells expressing the vector in peripheral blood. These results clearly show that the time of delivery of the vector is an important variable to take in account in an *in utero* delivery approach when the final objective is to target the hematopoietic progenitors located in the fetal liver.

Results obtained with the A2UCOE-eGFP-WPRE vector were confirmed by injection of a twin vector expressing luciferase, which allows vector expression to be monitored *in vivo* throughout the whole experimental time period, which lasted for up to 360 days (**Figure 5.13**). High expression of luciferase was observed in liver *in vivo* throughout the duration of the experiment. Post-mortem vector and protein expression analysis showed not only the feasibility of the *in utero* delivery system for BM engraftment but also long-term vector expression (up to 360 days with a luciferase reporter gene).

These reporter gene marking studies served as a precursor to our therapeutic study using a vector capable of expressing HBB to rescue a thalassaemia phenotype. However, it should be noted that this approach can also in principle be applied to other therapeutic studies, which aim to target HSCs. For instance, this approach can be used to target chromosome associated *Myelodysplastic Syndrome* (MDS), a group of diseases characterised by abnormal bone marrow cell production, which leads to symptoms of

anaemia, infection, excessive bleeding and bruising, and genetic defects such as *Fanconi Anaemia* resulting from mutation of one of a number of proteins responsible for DNA repair leading to development of leukaemia in adulthood.

7.2.3 *In utero* studies performed with a therapeutic *HBB*-based LV in WT mice

Following the promising reporter gene marking experiments described in Chapter 5, we continued our line of investigation by performing *in utero* injections with a therapeutic vector containing an *HBB* expression cassette. Due to problems in breeding and expanding the *th3* thalassaemic mouse colony as used in previous studies (Miccio *et al.*, 2008), we proceeded with initial experiments employing MF1 (wild type) mouse strain recipients.

Chapter 6 described the results from the delivery *in utero* of a *HBB* vector (GLOBE2) into wild type mice and showed sustained expression. Mice were analyzed at 5 and 8 months of age for *HBB* gene expression. As expected GLOBE2 vector was detected not only in peripheral blood and BM but also in liver and spleen.

To test *in vivo* different globin vectors and access their performance with regards to expression levels per copy number we used an improved, new version, GLOBE2, which is a development of the prototypical GLOBE vector, which has been shown to successfully rescue the thalassaemia phenotype in mice (Miccio *et al*., 2008) and human patient cells (Roselli *et al*., 2010). GLOBE2 was built based on results obtained with an HBB-eGFP LV showing that inclusion of the *HBB* transcription termination region (West and Proudfoot, 2009) results in up to a 4-fold increase in protein production both in cell lines *in vitro* and *in vivo* following *ex vivo* manipulation of HSC (Kao V *et al*., *unpublished results*).

The results obtained clearly show persistent GLOBE2 vector presence in peripheral blood and BM at both 150 and 250 days post-partum implying successful transduction of HSC in fetal liver following *in utero* injection. The results validate the initial hypothesis formulated in this study and show that it is possible to transduce HSC while in the fetal liver, before migration to the BM.

Expression levels obtained by mRNA analysis may be of a level capable of rescuing the thalassaemia phenotype and pave the way for a therapeutic experiment onto a thalassaemia mouse model system such as *th3* using the GLOBE2 vector. By the combination of the elements tested and others, the achievement of expression levels sufficient for therapeutic correction of the thalassaemia major phenotype can be anticipated.

The only report to date targeting *in utero* gene therapy for the haemoglobinopathies was performed with a *TNS9*-derived vector (**Figure 1.13),** which contains *HBA* under control of an extended *HBB* promoter, *HBB* proximal enhancers, and genomic fragments of the human *HBB* locus control region HS2, HS3 and HS4 (Xan *et al.*, 2007). In this study the authors claim that this LV can be an effective vehicle for delivering the *HBA* gene in cases of α-thalassaemia.

Intrauterine injection performed at E14.5 in 3 mice resulted in expression of *HBA* at birth and reached a peak of up to 20% in some mice at 3-4 months but the level of expression declined to 5% at approximately the 7th month. Very few erythroid colonies cultured from the bone marrow after decline of expression contained the transgene. The authors claim that the decline may be due to the loss of the transgene as a result of the inadequate transduction of HSC (Xan *et al*., 2007).

In our study, at 250 days (approximately 8 months) we observed variation of *HBB* transgene expression in the recipient mice obtained. This variation of expression is expected since *in vivo* delivery can be highly variable and, more importantly, the expression will be highly dependent on the vector integration sites. In our recipients injected *in utero* with the GLOBE2 vector, 3 showed a decrease in HBB levels per vector copy number, while 2 animals showed increased levels of HBB in peripheral blood during the course of the experiment (**Figure 6.4**). However, the degree of cell marking with GLOBE2 that we observe suggests transduction of HSCs, which is supported by the proviral copy number (average 0.371; **Figure 6.7**) and expression (**Figure 6.9**) found in BM upon sacrifice.

7.2.3.1 Injections *in utero* in β-thalassaemia *intermedia th3*$^{+/--}$ mice

It was part of the workplan of this study to conduct *in utero* injections as described here in a severe β-thalassaemia *intermedia* mouse model system, *th3*$^{+}$ (C57Bl/6 *Hbb*(th3/+)) with the GLOBE2 vector. As homozygous *th3/th3* mice die prenatally or perinatally, survival of these animals following *in utero* injection of GLOBE2 would be a clear indication of successful therapy.

Unfortunately, major difficulties in breeding the C57Bl/6 *Hbb*(th3/+) mice and time limitations impaired this phase of the work.

7.3 Final Conclusions

At present, it seems likely that future gene therapy vectors for the haemoglobinopathies will employ numerous methods to improve the efficacy and safety of gene therapy treatment used to replace blood transfusions. These may include the use of integration deficient vectors to decrease the risk of insertional mutagenesis and vector delivery at earlier developmental stages to avoid age onset malfunctions.

In this thesis we report a new therapeutic approach for the haemoglobinopathies. We describe the efficient transduction of HSC progenitors in fetal liver by *in utero* injection at 14-16dpc. Transduction of HSC is followed by successful BM engraftment and transgene levels in the systemic circulation.

Further studies are necessary to assess the feasibility of the *in utero* procedure in a human context, especially experiments in a murine model of thalassaemia followed by studies in larger animals such as sheep and primates. We envisage the development of new vector designs, capable of exploiting transduction of BM cells. These experiments have laid the foundation for new lines of investigation.

7.4 Future work

Leading on from the work presented here are several lines of research that would be interesting to pursue:

1. Addition of more globin LCR HS sites to the *HBB* cassette in the LV namely HSI and HSIV, in order to try and obtain higher levels of expression per vector copy.

2. For further experiments, higher MOIs should be used and since the volume administered in the *in utero* procedure cannot be increased and therefore it becomes essential to generate the *HBB* vector preparations at a higher titer.

3. Corrective gene therapy treatment would benefit from increased efficiency of transduction of HSC. It would therefore be very interesting to attempt to increase vector transduction of HSC by pseudotyping LVs with other envelope glycoproteins, such as feline leukemia virus (FeLV) RD114 and FeLV-C94 proteins. These FeLV envelope proteins have been shown to give rise to good human HSC transduction, and be less toxic to cells than VSV-G (Hanawa *et al.*, 2005).

4. It would be interesting to perform HBB transgene codon optimisation to attempt an improvement in protein production.

5. Finally, injection of larger animal models in which the fetal circulation can be accessed earlier and more readily and which is more similar to the human condition compared to mice.

REFERENCES

Ahn SH, Kim M, Buratowski S. (2004) Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell. Jan 16;13(1): 67-76.

Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L. et al. (2009) Gene therapy for immunodeficiency due to adenosine deaminase deficiency *N Engl J Med*., 360: 447-458.

Ali RR. (2012) Gene therapy for retinal dystrophies: twenty years in the making. *Hum Gene Ther*. 23: 337-339

Alexander SL, Linde-Zwirble WT, Werther W, Depperschmidt EE, Wilson LJ, Palanki R, Saroj N, Butterworth SL, Ianchulev T.(2007) Annual rates of arterial thromboembolic events in medicare neovascular age-related macular degeneration patients. Ophthalmology.Dec;114(12): 2174-2178.

Anderson WF, Blaese RM, Culver K. (1990) The ADA human gene therapy clinical protocol: Points to Consider response with clinical protocol, Jul Hum Gene Ther. Fall;1(3):331-362.

Antoniou M, Geraghty F, Hurst J, Grosveld F (1998) Efficient 3'-end formation of human beta-globin mRNA in vivo requires sequences within the last intron but occurs independently of the splicing reaction. Nucleic Acids Res. Feb 1;26(3): 721-729.

Antoniou M, Harland L, Mustoe T, Williams S, Holdstock J, Yague E, Mulcahy T, Griffiths M, Edwards S, Ioannou PA, Mountain A, Crombie R. (2003) Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics. Sep;82(3): 269-279.

Antoniou & Grosveld. (1999) *Genetic approaches to therapy for the haemoglobinopathies* at *Blood Cell Biochemistry,* Leslie J. Fairbairn & Nydia G. Testa (eds), 8: 219-234

Arumugam PI, Scholes J, Perelman N, Xia P, Yee JK, Malik P. (2007) Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. Mol Ther. Oct;15(10): 1863-1871.

Bainbridge JW, Ali RR (2008) Success in sight: The eyes have it! Ocular gene therapy trials for LCA look promising. Gene Ther. Sep;15(17): 1191-1192.

Aiuti A, Cattaneo F, Galimberti S, Benninghoff U, Cassani B, Callegaro L. et al. (2009) Gene therapy for immunodeficiency due to adenosine deaminase deficiency *N Engl J Med*., 360: 447-458.

Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, Viswanathan A, Holder GE, Stockman A, Tyler N, Petersen-Jones S, Bhattacharya SS, Thrasher AJ, Fitzke FW, Carter BJ, Rubin GS, Moore AT, Ali RR.(2008) Effect of gene therapy on visual function in Leber's congenital amaurosis. N Engl J Med.May 22;358(21): 2231-2239.

Bainbridge JW, Tan MH, Ali RR. (2006) Gene therapy progress and prospects: the eye. Gene Ther. Aug;13(16): 1191-1197.

Banasik MB, McCray PB Jr. (2009) Integrase-defective lentiviral vectors: progress and applications. Gene Ther. Feb;17(2): 150-157.

Banerjee A, Sammarco MC, Ditch S, Grabczyk E.(2009) A dual reporter approach to quantify defects in messenger RNA processing. Anal Biochem. Dec 15;395(2): 237-243.

Blaese RM, Culver KW, Chang L, Anderson WF, Mullen C, Nienhuis A, Carter C, Dunbar C, Leitman S, Berger M, (1993) Treatment of severe combined immunodeficiency disease (SCID) due to adenosine deaminase deficiency with CD34+ selected autologous peripheral blood cells transduced with a human ADA gene. Amendment to clinical research project, Project 90-C-195 Hum Gene Ther.Aug;4(4): 521-527.

Brendel C, Müller-Kuller U, Schultze-Strasser S, Stein S, Chen-Wichmann L, Krattenmacher A, Kunkel H, Dillmann A, Antoniou MN, Grez M. (2011) Physiological regulation of transgene expression by a lentiviral vector containing the A2UCOE linked to a myeloid promoter.Gene Ther. Nov 10. doi: 10.1038/gt.2011.167.

Brown KM, Gilmartin GM (2003) A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. Mol Cell.Dec;12(6): 1467-1476.

Bulger M, Groudine M. (1999) Looping versus linking: toward a model for long-distance gene activation. Genes Dev. Oct 1;13(19): 2465-2477.

Burns MK, Cooper KD. Cutaneous (1993) T-cell lymphoma associated with HIV infection. J Am Acad Dermatol. Sep;29(3): 394-399.

Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P.,(2002) Long-range chromatin regulatory interactions in vivo. Nat Genet.Dec;32(4): 623-626.

Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, Kutschera I, Vidaud M, Abel U, Dal-Cortivo L, Caccavelli L, Mahlaoui N, Kiermer V, Mittelstaedt D, Bellesme C, Lahlou N, Lefrère F, Blanche S, Audit M, Payen E, Leboulch P, l'Homme B, Bougnères P, Von Kalle C, Fischer A, Cavazzana-Calvo M, Aubourg P (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy, Science. Nov 6;326(5954): 818-823.

Cavazzana-Calvo M, Hacein-Bey S, de Saint Basile G, Gross F, Yvon E, Nusbaum P, Selz F, Hue C, Certain S, Casanova JL, Bousso P, Deist FL, Fischer A. (2000) Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science. Apr 28;288(5466): 669-672.

Cavazzana-Calvo M, Payen E, Negre O, Wang G, Hehir K, Fusil F, Down J, Denaro M, Brady T, Westerman K, Cavallesco R, Gillet-Legrand B, Caccavelli L, Sgarra R, Maouche-Chrétien L, Bernaudin F, Girot R, Dorazio R, Mulder GJ, Polack A, Bank A, Soulier J, Larghero J, Kabbara N, Dalle B, Gourmel B, Socie G, Chrétien S, Cartier N, Aubourg P, Fischer A, Cornetta K, Galacteros F, Beuzard Y, Gluckman E, Bushman F, Hacein-Bey-Abina S, Leboulch P. (2010) Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature. Sep 16;467(7313): 318-322.

Chang AH, Stephan MT, Sadelain M. (2006) Stem cell-derived erythroid cells mediate long-term systemic protein delivery. Nat Biotechnol. Aug;24(8): 1017-1021.

Chang AH, Stephan MT, Sadelain M. (2006) Stem cell-derived erythroid cells mediate long-term systemic protein delivery. Nat Biotechnol. Aug;24(8): 1017-1021.

Charneau P, Alizon M, Clavel F. (1992) A second origin of DNA plus-strand synthesis is required for optimal human immunodeficiency virus replication. J Virol. May;66(5): 2814-2820.

Chung DC, Traboulsi EI (2009) Leber congenital amaurosis: clinical correlations with genotypes, gene therapy trials update, and future directions J AAPOS. Dec;13(6): 587-589.

Coil DA, Miller AD. (2004) Phosphatidylserine is not the cell surface receptor for vesicular stomatitis virus. J Virol. Oct;78(20): 10920-10926.

Cone RD, Weber-Benarous A, Baorto D, Mulligan RC. (1987) Regulated expression of a complete human beta-globin gene encoded by a transmissible retrovirus vector. Mol Cell Biol. Feb;7(2): 887-897.

Connelly S, Manley JL. (1988) A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II.Genes Dev. Apr;2(4): 440-452.

Coutelle C, (2008) Why bother? Is in utero gene therapy worth the effort? Mol Ther.Feb;16(2): 219-220.

Coutelle C, Themis M, Waddington S, Gregory L, Nivsarkar M, Buckley S, Cook T, Rodeck C, Peebles D, David A. (2003) The hopes and fears of in utero gene therapy for genetic disease--a review, Placenta. Oct;24 Suppl B: S114-S121.

Coutelle C, Themis M, Waddington SN, Buckley SM, Gregory LG, Nivsarkar MS, David AL, Peebles D, Weisz B, Rodeck C., (2005) Gene therapy progress and prospects: fetal gene therapy--first proofs of concept--some adverse effects. Gene Ther. Nov;12(22): 1601-1607.

Cunningham MJ, Macklin EA, Neufeld EJ, Cohen AR (2004) Thalassemia Clinical Research Network, Complications of beta-thalassemia major in North America, Blood. Jul 1;104(1): 34-39

David AL, Peebles D. (2008) Gene therapy for the fetus: is there a future? Best Pract Res Clin Obstet Gynaecol. Feb;22(1): 203-218.

de Laat W, Grosveld F (2003) Spatial organization of gene expression: the active chromatin hub. Chromosome Res.;11(5): 447-459.

Dean A. (2006) On a chromosome far, far away: LCRs and gene expression.Trends Genet. 2006 Jan;22(1): 38-45. *Review*

deBoer E, Antoniou M, Mignotte V, Wall L, Grosveld F. (1988) The human beta-globin promoter; nuclear protein factors and erythroid specific induction of transcription. EMBO J. 1988 Dec 20;7(13): 4203-4012.

Dejneka NS, Kuroki AM, Fosnot J, Tang W, Tolentino MJ, Bennett J.(2004) Systemic rapamycin inhibits retinal and choroidal neovascularization in mice. Mol Vis. Dec 22;10: 964-972.

Delenda C. (2004) Lentiviral vectors: optimization of packaging, transduction and gene expression. J Gene Med. Feb;6 Suppl 1:S125-S138. *Review*

Demaison C, Parsley K, Brouns G, Scherr M, Battmer K, Kinnon C, Grez M, Thrasher AJ. (2002) High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Hum Gene Ther. May 1;13(7): 803-813.

DePolo NJ, Reed JD, Sheridan PL, Townsend K, Sauter SL, Jolly DJ, Dubensky TW Jr. (2000) VSV-G pseudotyped lentiviral vector particles produced in human cells are inactivated by human serum.Mol Ther. Sep;2(3): 218-222.

Desmaris N, Bosch A, Salaün C, Petit C, Prévost MC, Tordo N, Perrin P, Schwartz O, de Rocquigny H, Heard JM.(2001) Production and neurotropism of lentivirus vectors pseudotyped with lyssavirus envelope glycoproteins.Mol Ther. Aug;4(2): 149-156.

Dye MJ, Proudfoot NJ (2001) Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. Cell. Jun 1;105(5):669-681.

Dzierzak EA, Papayannopoulou T, Mulligan RC (1988) Lineage-specific expression of a human beta-globin gene in murine bone marrow transplant recipients reconstituted with retrovirus-transduced stem cells. Nature. Jan 7;331(6151): 35-41.

Einerhand MP, Antoniou M, Zolotukhin S, Muzyczka N, Berns KI, Grosveld F, Valerio D. (1995) Regulated high-level human beta-globin gene expression in erythroid cells following recombinant adeno-associated virus-mediated gene transferGene Ther. Jul;2(5): 336-343

Ellis J. (2005) Silencing and variegation of gammaretrovirus and lentivirus vectors. Hum Gene Ther. Nov;16(11): 1241-1246.

Emerman M, Temin HM.(1984) High-frequency deletion in recovered retrovirus vectors containing exogenous DNA with promoters. J Virol. Apr;50(1): 42-49.

Engelman A. (1999) In vivo analysis of retroviral integrase structure and function.Adv Virus Res.;52: 411-426. *Review*

Escors D, Breckpot K. (2010) Lentiviral vectors in gene therapy: their current status and future potential. Arch Immunol Ther Exp (Warsz). Apr;58(2): 107-119.

Feng YQ, Warin R, Li T, Olivier E, Besse A, Lobell A, Fu H, Lin CM, Aladjem MI, Bouhassira EE. (2005) The human beta-globin locus control region can silence as well as activate gene expression. Mol Cell Biol. May;25(10): 3864-3874.

Flotte TR.(2008) Gene therapy: the first two decades and the current state-of-the-art. J Cell Physiol.Nov;213(2): 301-305. *Review*

Follenzi A, Ailles LE, Bakovic S, Geuna M, Naldini L.(2000) Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. Nat Genet. Jun;25(2): 217-222.

Fraser P, Grosveld F, (1998) Locus control regions, chromatin activation and transcription. Curr Opin Cell Biol. Jun;10(3): 361-365.

Furger A, Monks J, Proudfoot NJ.(2001) The retroviruses human immunodeficiency virus type 1 and Moloney murine leukemia virus adopt radically different strategies to regulate promoter-proximal polyadenylation. J Virol. Dec;75(23): 11735-11746.

Fusco A, Fedele M. (2007) Roles of HMGA proteins in cancer. Nat Rev Cancer. Dec;7(12): 899-910.

Galanello R, Origa R (2010).Beta-thalassemia. Orphanet J Rare Dis. May 21; 5-11.

Gaspar HB, Parsley KL, Howe S, King D, Gilmour KC, Sinclair J, Brouns G, Schmidt M, Von Kalle C, Barington T, Jakobsen MA, Christensen HO, Al Ghonaium A, White HN, Smith JL, Levinsky RJ, Ali RR, Kinnon C, Thrasher AJ. (2004) Gene therapy of X-linked

severe combined immunodeficiency by use of a pseudotyped gammaretroviral vector. Lancet. Dec 18-31;364(9452):2181-2187.

Gaziev J, Lucarelli G. (2010) Allogeneic cellular gene therapy for hemoglobinopathies. Hematol Oncol Clin North Am. Dec;24(6): 1145-1163.

Gerton JL, Herschlag D, Brown PO, (1999) Stereospecificity of reactions catalyzed by HIV-1 integrase. J Biol Chem. Nov 19;274(47): 33480-33487.

Gregorevic P, Blankinship MJ, Chamberlain JS.(2004) Viral vectors for gene transfer to striated muscle.Curr Opin Mol Ther. Oct;6(5): 491-498. *Review*

Hacein-Bey-Abina S, Le Deist F, Carlier F, Bouneaud C, Hue C, De Villartay JP, Thrasher AJ, Wulffraat N, Sorensen R, Dupuis-Girod S, Fischer A, Davies EG, Kuis W, Leiva L, Cavazzana-Calvo M. (2002) Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N Engl J Med. Apr 18;346(16): 1185-1193.

Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, Sorensen R, Forster A, Fraser P, Cohen JI, de Saint Basile G, Alexander I, Wintergerst U, Frebourg T, Aurias A, Stoppa-Lyonnet D, Romana S, Radford-Weiss I, Gross F, Valensi F, Delabesse E, Macintyre E, Sigaux F, Soulier J, Leiva LE, Wissler M, Prinz C, Rabbitts TH, Le Deist F, Fischer A, Cavazzana-Calvo M. (2003a) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1.Science. Oct 17;302(5644):415-419.

Hacein-Bey-Abina S, von Kalle C, Schmidt M, Le Deist F, Wulffraat N, McIntyre E, Radford I, Villeval JL, Fraser CC, Cavazzana-Calvo M, Fischer A.(2003b) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med. Jan 16;348(3):255-256.

Hagedorn C, Wong SP, Harbottle R and Lipps HJ (2011) Scaffold/matrix attached region-based nonviral episomal vectors. Hum Gene Ther., 22: 915-923.

Han XD, Lin C, Chang J, Sadelain M, Kan YW.(2007) Fetal gene therapy of alpha-thalassemia in a mouse model. Proc Natl Acad Sci U S A. May 22;104(21):9007-9011

Hanawa H, Hargrove PW, Kepes S, Srivastava DK, Nienhuis AW, Persons DA (2004) Extended beta-globin locus control region elements promote consistent therapeutic

expression of a gamma-globin lentiviral vector in murine beta-thalassemia. Blood. Oct 15;104(8): 2281-2290.

Hanawa H, Kelly PF, Nathwani AC(2002) Comparison of various envelope proteins for their ability to pseudotype lentiviral vectors and transduce primitive hematopoietic cells from human blood. Mol Ther;5: 242–251.

Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, Miller W. (1997) Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights, Gene. Dec 31;205(1-2): 73-94.

Harteveld CL, Higgs DR. (2010) Alpha-thalassaemia. Orphanet J Rare Dis. May 28; 5-13.

Hatziioannou T, Delahaye E, Martin F, Russell SJ, Cosset FL.(1999) Retroviral display of functional binding domains fused to the amino terminus of influenza hemagglutinin. Hum Gene Ther. 1999 Jun 10;10(9): 1533-1544.

Hatzoglou M, Lamers W, Bosch F, Wynshaw-Boris A, Clapp DW, Hanson RW (1990) Hepatic gene transfer in animals using retroviruses containing the promoter from the gene for phosphoenolpyruvate carboxykinase. J Biol Chem. Oct 5;265(28):17285-17293.

Hauswirth WW, Aleman TS, Kaushal S, Cideciyan AV, Schwartz SB, Wang L, Conlon TJ, Boye SL, Flotte TR, Byrne BJ, Jacobson SG. (2008) Treatment of leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. Hum Gene Ther. Oct;19(10):979-990.

Hauswirth WW, Aleman TS, Kaushal S, Cideciyan AV, Schwartz SB, Wang L, Conlon TJ, Boye SL, Flotte TR, Byrne BJ, Jacobson SG. (2008) Treatment of leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. Hum Gene Ther. Oct;19(10): 979-990.

Heilbronn R, Weger S. (2010) Viral vectors for gene transfer: current status of gene therapeutics. Handb Exp Pharmacol.;(197): 143-170.

Higgs DR, Weatherall DJ. (2009) The alpha thalassaemias, Cell Mol Life Sci. Apr;66(7): 1154-1162.

Howe SJ, Mansour MR, Schwarzwaelder K, Bartholomae C, Hubank M, Kempski H, Brugman MH, Pike-Overzet K, Chatters SJ, de Ridder D, Gilmour KC, Adams S, Thornhill SI, Parsley KL, Staal FJ, Gale RE, Linch DC, Bayford J, Brown L, Quaye M, Kinnon C, Ancliff P, Webb DK, Schmidt M, von Kalle C, Gaspar HB, Thrasher AJ.(2008) Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. J Clin Invest.Sep;118(9): 3143-3150.

Ikeda Y, Goto Y, Yonemitsu Y, Miyazaki M, Sakamoto T, Ishibashi T, Tabata T, Ueda Y, Hasegawa M, Tobimatsu S, Sueishi K.(2003) Simian immunodeficiency virus-based lentivirus vector for retinal gene transfer: a preclinical safety study in adult rats.Gene Ther. 2003 Jul;10(14): 1161-1169.

Imren S, Fabry ME, Westerman KA, Pawliuk R, Tang P, Rosten PM, Nagel RL, Leboulch P, Eaves CJ, Humphries RK.(2004)High-level beta-globin expression and preferred intragenic integration after lentiviral transduction of human cord blood stem cells. J Clin Invest. Oct;114(7): 953-962.

Kappes JC, Wu X (2001) Safety considerations in vector development. Somat Cell Mol Genet. Nov;26(1-6): 147-158.

Katsantoni EZ, Anghelescu NE, Rottier R, Moerland M, Antoniou M, de Crom R, Grosveld F, Strouboulis J. (2007) Ubiquitous expression of the rtTA2S-M2 inducible system in transgenic mice driven by the human hnRNPA2B1/CBX3 CpG island.BMC Dev Biol.Sep 27;7:108-114.

Katz RA, Skalka AM. (1994) The retroviral enzymes. Annu Rev Biochem. ;63: 133-173.

Kazazian HH Jr. (1999) An estimated frequency of endogenous insertional mutations in humans. Nat Genet. Jun; 22(2): 130.

Kikuchi K, Kondo M.(2008) Developmental switch of mouse hematopoietic stem cells from fetal to adult type occurs in bone marrow after birth.Proc Natl Acad Sci U S A. Nov 21;103(47): 17852-17857.

Kohn DB, Hershfield MS, Carbonaro D, Shigeoka A, Brooks J, Smogorzewska EM, Barsky LW, Chan R, Burotto F, Annett G, Nolta JA, Crooks G, Kapoor N, Elder M, Wara D, Bowen T, Madsen E, Snyder FF, Bastian J, Muul L, Blaese RM, Weinberg K, Parkman R.(1998) T

lymphocytes with a normal ADA gene accumulate after transplantation of transduced autologous umbilical cord blood CD34+ cells in ADA-deficient SCID neonates. Nat Med. Jul;4(7):775-780.

Kotin RM, Siniscalco M, Samulski RJ, Zhu XD, Hunter L, Laughlin CA, McLaughlin S, Muzyczka N, Rocchi M, Berns KI.(1990) Site-specific integration by adeno-associated virus. Proc Natl Acad Sci U S A. Mar;87(6): 2211-2215.

Kuate S, Stefanou D, Hoffmann D, Wildner O, Uberla K. (2004) Production of lentiviral vectors by transient expression of minimal packaging genes from recombinant adenoviruses. J Gene Med.Nov;6(11): 1197-1205.

Larson JE, Morrow SL, Happel L, Sharp JF, Cohen JC (1997) Reversal of cystic fibrosis phenotype in mice by gene therapy in utero. Lancet. Mar 1;349(9052): 619-620.

Lee CC, Jimenez DF, Kohn DB, Tarantal AF (2005) Fetal gene transfer using lentiviral vectors and the potential for germ cell transduction in rhesus monkeys (Macaca mulatta). Hum Gene Ther. Apr;16(4): 417-425.

Levasseur DN, Ryan TM, Pawlik KM, Townes TM (2003) Correction of a mouse model of sickle cell disease: lentiviral/antisickling beta-globin gene transduction of unmobilized, purified hematopoietic stem cells. Blood. Dec 15;102(13): 4312-4319.

Levings PP, Bungert J. (2002) The human beta-globin locus control region. Eur J Biochem. Mar;269(6):1589-1599.

Li Q, Harju S, Peterson KR, (1999) Locus control regions: coming of age at a decade plus. Trends Genet. Oct;15(10):403-408.

Li Q, Peterson KR, Fang X, Stamatoyannopoulos G, (2002) Locus control regions. Blood. Nov 1;100(9): 3077-3086.

Li Z, Paulin D. (1993) Different factors interact with myoblast-specific and myotube-specific enhancer regions of the human desmin gene.J Biol Chem. May 15;268(14): 10403-10415.

Lindahl Allen M, Antoniou M.(2007) Correlation of DNA methylation with histone modifications across the HNRPA2B1-CBX3 ubiquitously-acting chromatin open element (UCOE). Epigenetics. Oct-Dec;2(4): 227-236.

Logan J, Falck-Pedersen E, Darnell JE Jr, Shenk T. (1987) A poly(A) addition site and a downstream termination region are required for efficient cessation of transcription by RNA polymerase II in the mouse beta maj-globin gene. Proc Natl Acad Sci U S A. Dec;84(23): 8306-8310.

Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD, Holmes MC, Naldini L.(2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery.Nat Biotechnol. Nov;25(11):1298-1306.

Luo W, Zhou X, Tian X, Ren X, Zheng M, Gu K, He G. (2006) Enhancement of ultrasound contrast agent in high-intensity focused ultrasound ablation.Adv Ther. Nov-Dec;23(6): 861-868. *Review*

Lutzko C, Senadheera D, Skelton D, Petersen D, Kohn DB. (2003) Lentivirus vectors incorporating the immunoglobulin heavy chain enhancer and matrix attachment regions provide position-independent expression in B lymphocytes. J Virol. Jul;77(13): 7341-7351.

Madigan C, Malik P.(2006) Pathophysiology and therapy for haemoglobinopathies. Part I: sickle cell disease. Expert Rev Mol Med. Apr 28;8(9): 1-23.

Maguire AM, Simonelli F, Pierce EA, Pugh EN Jr, Mingozzi F, Bennicelli J, Banfi S, Marshall KA, Testa F, Surace EM, Rossi S, Lyubarsky A, Arruda VR, Konkle B, Stone E, Sun J, Jacobs J, Dell'Osso L, Hertle R, Ma JX, Redmond TM, Zhu X, Hauck B, Zelenaia O, Shindler KS, Maguire MG, Wright JF, Volpe NJ, McDonnell JW, Auricchio A, High KA, Bennett J (2008) Safety and efficacy of gene transfer for Leber's congenital amaurosis. N Engl J Med. May 22;358(21): 2240-2248.

Maguire MG, Alexander J, Fine SL; (2008) Complications of Age-related Macular Degeneration Prevention Trial (CAPT) Research Group. Characteristics of choroidal neovascularization in the complications of age-related macular degeneration prevention trial. Ophthalmology. Sep;115(9): 1468-1173

May C, Rivella S, Callegari J, Heller G, Gaensler KM, Luzzatto L, Sadelain M. (2000) Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. Nature. Jul 6;406(6791): 82-86.

May C, Rivella S, Chadburn A, Sadelain M. (2002) Successful treatment of murine beta-thalassemia intermedia by transfer of the human beta-globin gene. Blood. Mar 15;99(6): 1902-1908.

Mazarakis ND, Azzouz M, Rohll JB, Ellard FM, Wilkes FJ, Olsen AL, Carter EE, Barber RD, Baban DF, Kingsman SM, Kingsman AJ, O'Malley K, Mitrophanous KA.(2001) Rabies virus glycoprotein pseudotyping of lentiviral vectors enables retrograde axonal transport and access to the nervous system after peripheral delivery.Hum Mol Genet.Sep 15;10(19): 2109-2121.

McGaughey GB, Barbato G, Bianchi E, Freidinger RM, Garsky VM, Hurni WM, Joyce JG, Liang X, Miller MD, Pessi A, Shiver JW, Bogusky MJ.(2004) Progress towards the development of a HIV-1 gp41-directed vaccine. Curr HIV Res. Apr;2(2): 193-204.

Meerpohl JJ, Antes G, Rücker G, Fleeman N, Motschall E, Niemeyer CM, Bassler D.(2010) Deferasirox for managing iron overload in people with thalassaemia. Cochrane Database Syst Rev. Feb 15;2: CD007476.

Miccio A, Cesari R, Lotti F, Rossi C, Sanvito F, Ponzoni M, Routledge SJ, Chow CM, Antoniou MN, Ferrari G (2008) In vivo selection of genetically modified erythroblastic progenitors leads to long-term correction of beta-thalassemia. Proc Natl Acad Sci U S A. Jul 29;105(30):10547-10552.

Michallet M, Tanguy ML, Socié G, Thiébaut A, Belhabri A, Milpied N, Reiffers J, Kuentz M, Cahn JY, Blaise D, Demeocq F, Jouet JP, Michallet AS, Ifrah N, Vilmer E, Molina L, Michel G, Lioure B, Cavazzana-Calvo M, Pico JL, Sadoun A, Guyotat D, Attal M, Curé H, Bordigoni P, Sutton L, Buzyn-Veil A, Tilly M, Keoirruer N, Feguex N. (2000) Second allogeneic haematopoietic stem cell transplantation in relapsed acute and chronic leukaemias for patients who underwent a first allogeneic bone marrow transplantation: a survey of the Société Française de Greffe de moelle (SFGM). Br J Haematol. Feb;108(2): 400-407.

Millevoi S, Geraghty F, Idowu B, Tam JL, Antoniou M, Vagner SA (2002) novel function for the U2AF 65 splicing factor in promoting pre-mRNA 3'-end processing. EMBO Rep. Sep;3(9): 869-874.

Milsom MD, Williams DA., (2010) Gaining the hard yard: pre-clinical evaluation of lentiviral-mediated gene therapy for the treatment of beta-thalassemia, EMBO Mol Med. Aug;2(8): 291-293.

Miyoshi H, Blömer U, Takahashi M, Gage FH, Verma IM. (1998) Development of a self-inactivating lentivirus vector. J Virol. Oct;72(10): 8150-8157.

Mochizuki H, Schwartz JP, Tanaka K, Brady RO, Reiser J.(1998) High-titer human immunodeficiency virus type 1-based vector systems for gene delivery into nondividing cells. J Virol. 1998 Nov;72(11): 8873-8883.

Modlich U, Bohne J, Schmidt M, von Kalle C, Knöss S, Schambach A, Baum C.(2006) Cell-culture assays reveal the importance of retroviral vector design for insertional genotoxicity. Blood. Oct 15;108(8): 2545-2553.

Moreno-Carranza B, Gentsch M, Stein S, Schambach A, Santilli G, Rudolf E, Ryser MF, Haria S, Thrasher AJ, Baum C, Brenner S, Grez M (2009) Transgene optimization significantly improves SIN vector titers, gp91phox expression and reconstitution of superoxide production in X-CGD cells. Gene Ther.Jan;16(1): 111-118.

Morizono K, Chen IS.(2005) Targeted gene delivery by intravenous injection of retroviral vectors.Cell Cycle. Jul;4(7): 854-856.

Murthy KG, Manley JL.(1995) The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. Genes Dev. Nov 1;9(21): 2672-2683.

Nair AR, Jinger X, Hermiston TW. (2011) Effect of different UCOE-promoter combinations in creation of engineered cell lines for the production of Factor VIII. BMC Res Notes. Jun 10;4: 178.

Naldini L, Blömer U, Gage FH, Trono D, Verma IM. (1996) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. Proc Natl Acad Sci U S A. Oct 15;93(21): 11382-11388.

Noble R, Rodeck CH, (2008) Ethical considerations of fetal therapy. Best Pract Res Clin Obstet Gynaecol. Feb;22(1): 219-231.

Ott M.G., Schmidt M., Schwarzwaelder K., Stein S., Siler U., Koehl U., Glimm H., Kühlcke K., Schilz A., Kunkel H., Naundorf S., Brinkmann A., Deichmann A., Fischer M., Ball C., Pilz I., Dunbar C., Du Y., Jenkins N.A., Copeland N.G., Lüthi U., Hassan M., Thrasher A.J., Hoelzer D., von Kalle C., Seger R.; Grez M., (2006) Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of M DS1-EVI1, PRDM16 or SETBP1, Nat. Med. 12 , 401–409.

Palstra RJ, de Laat W, Grosveld F. (2008) Beta-globin regulation and long-range interactions. Adv Genet.;61: 107-142.

Pawliuk R, Westerman KA, Fabry ME, Payen E, Tighe R, Bouhassira EE, Acharya SA, Ellis J, London IM, Eaves CJ, Humphries RK, Beuzard Y, Nagel RL, Leboulch P.(2001) Correction of sickle cell disease in transgenic mouse models by gene therapy Science. Dec 14;294(5550): 2368-2371.

Persons DA. (2010) Gene therapy: Targeting β-thalassaemia. Nature. Sep 16;467(7313): 277.

Pichlmair A, Diebold SS, Gschmeissner S, Takeuchi Y, Ikeda Y, Collins MK, Reis e Sousa C. (2007) Tubulovesicular structures within vesicular stomatitis virus G protein-pseudotyped lentiviral vector preparations carry DNA and stimulate antiviral responses via Toll-like receptor 9.J Virol Jan;81(2): 539-547.

Pluta K, Kacprzak MM (2009) Use of HIV and gene transfer vector. Acta Biochim Pol. ;56(4): 531-595.

Poeschla E, Corbeau P, Wong-Staal F. (1996) Development of HIV vectors for anti-HIV gene therapy. Proc Natl Acad Sci U S A. Oct 15;93(21): 11395-11399.

Porada CD, Tran N, Eglitis M, Moen RC, Troutman L, Flake AW, Zhao Y, Anderson WF, Zanjani ED (1998) In utero gene therapy: transfer and long-term expression of the bacterial neo(r) gene in sheep after direct injection of retroviral vectors into preimmune fetuses., Hum Gene Ther. Jul 20;9(11): 1571-1585.

Puthenveetil, G. et al. (2004) Successful correction of the human beta-thalassemia major phenotype using a lentiviral vector. *Blood* 104(12), 3445–3453.

Quek L, Thein SL. (2007) Molecular therapies in beta-thalassaemia. Br J Haematol. 2007 Feb;136(3):353-365.

Raguz S, Hobbs C, Yagüe E, Ioannou PA, Walsh FS, Antoniou M. (1998) Muscle-specific locus control region activity associated with the human desmin gene. Dev Biol. 1998 Sep 1;201(1): 26-42.

Rahim AA, Wong AM, Howe SJ, Buckley SM, Acosta-Saltos AD, Elston KE, Ward NJ, Philpott NJ, Cooper JD, Anderson PN, Waddington SN, Thrasher AJ, Raivich G. (2009) Efficient gene delivery to the adult and fetal CNS using pseudotyped non-integrating lentiviral vectors.Gene Ther. Apr;16(4): 509-520.

Rattray AJ, Champoux JJ. (1989) Plus-strand priming by Moloney murine leukemia virus. The sequence features important for cleavage by RNase H. J Mol Biol. Aug 5;208(3): 445-456.

Reiser J, Harmison G, Kluepfel-Stahl S, Brady RO, Karlsson S, Schubert M. (1996) Transduction of nondividing cells using pseudotyped defective high-titer HIV type 1 particles. Proc Natl Acad Sci U S A.Dec 24;93(26): 15266-15271.

Rivella S, May C, Chadburn A, Rivière I, Sadelain M. (2003) A novel murine model of Cooley anemia and its rescue by lentiviral-mediated human beta-globin gene transfer Blood. Apr 15;101(8): 2932-2939.

Rivella S, Sadelain M. (1998) Genetic treatment of severe hemoglobinopathies: the combat against transgene variegation and transgene silencing. Semin Hematol. Apr;35(2): 112-125.

Roselli EA, Mezzadra R, Frittoli MC, Maruggi G, Biral E, Mavilio F, Mastropietro F, Amato A, Tonon G, Refaldi C, Cappellini MD, Andreani M, Lucarelli G, Roncarolo MG, Marktel S, Ferrari G, (2010) Correction of beta-thalassemia major by gene transfer in haematopoietic progenitors of pediatric patients. EMBO Mol Med. Aug;2(8): 315-328.

Rucker M, Fraites TJ Jr, Porvasnik SL, Lewis MA, Zolotukhin I, Cloutier DA, Byrne BJ.(2004) Rescue of enzyme deficiency in embryonic diaphragm in a mouse model of metabolic myopathy: Pompe disease. Development. Jun;131(12): 3007-3019.

Rund D, Rachmilewitz E., (2005) Beta-thalassemia. N Engl J Med. Sep 15;353(11): 1135-1146.

Ruscetti SK.(1999) Deregulation of erythropoiesis by the Friend spleen focus-forming virus.Int J Biochem Cell Biol. Oct;31(10): 1089-1109. *Review*

Ryser MF, Roesler J, Gentsch M, Brenner S, (2007) Gene therapy for chronic granulomatous disease. Expert Opin Biol Ther. Dec;7(12): 1799-1809.

Sadelain M, Lisowski L, Samakoglu S, Rivella S, May C, Riviere I. (2005) Progress toward the genetic treatment of the beta-thalassemias. Ann N Y Acad Sci. ;1054: 78-91.

Sadelain M, Wang CH, Antoniou M, Grosveld F, Mulligan RC. (1995) Generation of a high-titer retroviral vector capable of expressing high levels of the human beta-globin gene. Proc Natl Acad Sci U S A. Jul 18;92(15): 6728-6732.

Sadelain M. (2006) Recent advances in globin gene transfer for the treatment of beta-thalassemia and sickle cell anemia. Curr Opin Hematol. May;13(3): 142-148.

Salva MZ, Himeda CL, Tai PW, Nishiuchi E, Gregorevic P, Allen JM, Finn EE, Nguyen QG, Blankinship MJ, Meuse L, Chamberlain JS, Hauschka SD. (2007) Design of tissue-specific regulatory cassettes for high-level rAAV-mediated expression in skeletal and cardiac muscle. Mol Ther. Feb;15(2): 320-329.

Schachtner S, Buck C, Bergelson J, Baldwin H. (1999) Temporally regulated expression patterns following in utero adenovirus-mediated gene transfer. Gene Ther Jul;6(7): 1249-1257.

Schlegel R, Willingham MC, Pastan IH. (1982) Saturable binding sites for vesicular stomatitis virus on the surface of Vero cells.J Virol. 1982 Sep;43(3): 871-875.

Seppen J, van der Rijt R, Looije N, van Til NP, Lamers WH, Oude Elferink RP (2003) Long-term correction of bilirubin UDPglucuronyltransferase deficiency in rats by in utero lentiviral gene transfer. Mol Ther. Oct;8(4): 593-599.

Shank PR, Hughes SH, Kung HJ, Majors JE, Quintrell N, Guntaka RV, Bishop JM, Varmus HE. (1978) Mapping unintegrated avian sarcoma virus DNA: termini of linear DNA bear 300 nucleotides present once or twice in two species of circular DNA. Cell. 1978 Dec;15(4): 1383-1395.

Shaw SW, Bollini S, Nader KA, Gastadello A, Mehta V, Filppi E, Cananzi M, Gaspar HB, Qasim W, De Coppi P, David AL. (2011) Autologous transplantation of amniotic fluid-derived mesenchymal stem cells into sheep fetuses. Cell Transplant. 2011;20(7): 1015-1031.

Simonelli F, Maguire AM, Testa F, Pierce EA, Mingozzi F, Bennicelli JL, Rossi S, Marshall K, Banfi S, Surace EM, Sun J, Redmond TM, Zhu X, Shindler KS, Ying GS, Ziviello C, Acerra C, Wright JF, McDonnell JW, High KA, Bennett J, Auricchio A. (2010) Gene therapy for Leber's congenital amaurosis is safe and effective through 1.5 years after vector administration. Mol Ther. Mar;18(3): 643-650.

Staunstrup NH, Moldt B, Mátés L, Villesen P, Jakobsen M, Ivics Z, Izsvák Z, Mikkelsen JG (2009). Hybrid lentivirus-transposon vectors with a random integration profile in human cells. Mol Ther. Jul;17(7): 1205-1214.

Stein S, Ott MG, Schultze-Strasser S, Jauch A, Burwinkel B, Kinner A, Schmidt M, Krämer A, Schwäble J, Glimm H, Koehl U, Preiss C, Ball C, Martin H, Göhring G, Schwarzwaelder K, Hofmann WK, Karakaya K, Tchatchou S, Yang R, Reinecke P, Kühlcke K, Schlegelberger B, Thrasher AJ, Hoelzer D, Seger R, von Kalle C, Grez M. (2010) Genomic instability and myelodysplasia with monosomy 7 consequent to EVI1 activation after gene therapy for chronic granulomatous disease. Nat Med. Feb;16(2): 198-204.

Strathdee SA, Stachowiak JA, Todd CS, Al-Delaimy WK, Wiebel W, Hankins C, Patterson TL. (2006) Complex emergencies, HIV, and substance use: no "big easy" solution. Subst Use Misuse. ;41(10-12): 1637-1651.

Surbek DV, Holzgreve W, Nicolaides KH (2001) Haematopoietic stem cell transplantation and gene therapy in the fetus: ready for clinical use?.Hum Reprod Update. Jan-Feb;7(1): 85-91.

Talbot D, Grosveld F, (1991) The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J. Jun;10(6): 1391-1398.

Talbot GE, Waddington SN, Bales O, Tchen RC, Antoniou MN. (2010) Desmin-regulated lentiviral vectors for skeletal muscle gene transfer. Mol Ther. 2010 Mar;18(3): 601-608.

Tam JL, Triantaphyllopoulos K, Todd H, Raguz S, de Wit T, Morgan JE, Partridge TA, Makrinou E, Grosveld F, Antoniou M.(2006) The human desmin locus: gene organization and LCR-mediated transcriptional control.Genomics. Jun;87(6): 733-746.

Tarantal AF, Lee CC, Jimenez DF, Cherry SR.(2006) Fetal gene transfer using lentiviral vectors: in vivo detection of gene expression by microPET and optical imaging in fetal and infant monkeys. Hum Gene Ther. Dec;17(12): 1254-1261.

Termo Scientific Tech notes, Josh Haimes & Melissa Kelley, 2010

Themis M, Waddington SN, Schmidt M, von Kalle C, Wang Y, Al-Allaf F, Gregory LG, Nivsarkar M, Themis M, Holder MV, Buckley SM, Dighe N, Ruthe AT, Mistry A, Bigger B, Rahim A, Nguyen TH, Trono D, Thrasher AJ, Coutelle C.(2005) Oncogenesis following delivery of a nonprimate lentiviral gene therapy vector to fetal and neonatal mice. Mol Ther. Oct;12(4): 763-771.

Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell. Dec;10(6):1453-1465.

Trask RV, Strauss AW, Billadello JJ. (1988) Developmental regulation and tissue-specific expression of the human muscle creatine kinase gene. J Biol Chem. Nov 15;263(32): 17142-17149.

Troeger C, Surbek D, Schöberlein A, Schatt S, Dudler L, Hahn S, Holzgreve W, (2007) In utero haematopoietic stem cell transplantation. Experiences in mice, sheep and humans Swiss Med Wkly. Mar 2;137 Suppl 155: 14S-19S.

Verma IM, Somia N. (1997) Gene therapy -- promises, problems and prospects. Nature. Sep 18;389(6648): 239-242.

Vink CA, Gaspar HB, Gabriel R, Schmidt M, McIvor RS, Thrasher AJ, Qasim W. (2009) Sleeping beauty transposition from nonintegrating lentivirus.Mol Ther.Jul;17(7): 1197-1204.

von Kalle C, Baum C, Williams DA. (2004) Lenti in red: progress in gene therapy for human hemoglobinopathies. J Clin Invest. Oct;114(7): 889-891.

Waddington SN, Kennea NL, Buckley SM, Gregory LG, Themis M, Coutelle C.(2004) Fetal and neonatal gene therapy: benefits and pitfalls. Gene Ther. Oct;11 Suppl 1: S92-S97.

Waddington SN, Kramer MG, Hernandez-Alcoceba R, Buckley SM, Themis M, Coutelle C, Prieto J. (2005) In utero gene therapy: current challenges and perspectives. Mol Ther. May;11(5): 661-676.

Waddington SN, Nivsarkar MS, Mistry AR, Buckley SM, Kemball-Cook G, Mosley KL, Mitrophanous K, Radcliffe P, Holder MV, Brittan M, Georgiadis A, Al-Allaf F, Bigger BW, Gregory LG, Cook HT, Ali RR, Thrasher A, Tuddenham EG, Themis M, Coutelle (2004) Permanent phenotypic correction of hemophilia B in immunocompetent mice by prenatal gene therapy. Blood. Nov 1;104(9): 2714-2721.

Wagner AM, Schoeberlein A, Surbek D.(2009) Fetal gene therapy: opportunities and risks. Adv Drug Deliv Rev. Aug 10;61(10): 813-821.

Watanabe S, Temin HM. (1989) Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5' long terminal repeat and the start of the gag gene. Proc Natl Acad Sci U S A. Oct;79(19): 5986-5990.

Weatherall DJ, Clegg JB.(2001) Inherited haemoglobin disorders: an increasing global health problem.Bull World Health Organ;79(8): 704-712.

Weatherall DJ., (2010) Thalassemia as a global health problem: recent progress toward its control in the developing countries. Ann N Y Acad Sci. Aug;1202: 17-23.

West S, Proudfoot NJ, Dye MJ. (2008) Molecular dissection of mammalian RNA polymerase II transcriptional termination. Mol Cell. Mar 14;29(5): 600-610.

West S, Proudfoot NJ.(2009) Transcriptional termination enhances protein expression in human cells. Mol Cell. Feb 13;33(3): 354-364.

White MT.(2009) Prenatal diagnosis: spiritual challenges in genetic testing and counseling. Bull Park Ridge Cent. Jan-Feb;(13): 9-10.

Williams DA.(2006) New AAV serotypes may broaden the therapeutic pipeline to human gene therapy.Mol Ther. Jan;13(1): 1-2.

Winer J, Jung CK, Shackel I, Williams PM. (1999) Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes in vitro. Anal Biochem. May 15;270(1): 41-49.

Yáñez-Muñoz RJ, Balaggan KS, MacNeil A, Howe SJ, Schmidt M, Smith AJ, Buch P, MacLaren RE, Anderson PN, Barker SE, Duran Y, Bartholomae C, von Kalle C, Heckenlively JR, Kinnon C, Ali RR, Thrasher AJ.(2006) Effective gene therapy with nonintegrating lentiviral vectors. Nat Med. Mar;12(3): 348-353.

Ye, L. et al. (2009) Induced pluripotent stem cells offer new approach to therapy in thalassemia and sickle cell anemia and option in prenatal diagnosis in genetic diseases. *Proc. Natl. Acad. Sci. U. S. A.* 106(24); 9826-9830.

Yee TT, Pasi KJ, Lilley PA, Lee CA.(1987) Factor VIII inhibitors in haemophiliacs: a single-centre experience over 34 years, 1964-97.Br J Haematol. 1999 Mar;104(4): 909-914.

Yu JC, Nash MA, Santiago C, Marzluff WF.(1986) Structure and expression of a second sea urchin U1 RNA gene repeatNucleic Acids Res. Dec 22;14(24): 9977-9988.

Zaiss AK, Son S, Chang LJ. (2002) RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy.J Virol. Jul;76(14): 7209-7219.

Zennou V, Petit C, Guetard D, Nerhbass U, Montagnier L, Charneau P.(2000) HIV-1 genome nuclear import is mediated by a central DNA flap.Cell. Apr 14;101(2): 173-185.

Zhang F, Thornhill SI, Howe SJ, Ulaganathan M, Schambach A, Sinclair J, Kinnon C, Gaspar HB, Antoniou M, Thrasher AJ (2007) Lentiviral vectors containing an enhancer-less ubiquitously acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in hematopoietic cells. Blood. Sep 1;110(5): 1448-1457.

Zhang X, Hamada J, Nishimoto A, Takahashi Y, Murai T, Tada M, Moriuchi T. (2007b) HOXC6 and HOXC11 increase transcription of S100beta gene in BrdU-induced in vitro differentiation of GOTO neuroblastoma cells into Schwannian cells. J Cell Mol Med. Mar-Apr;11(2): 299-306.

Zhang F, Frost AR, Blundell MP, Bales O, Antoniou MN, Thrasher AJ.(2010) A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors.Mol Ther. 2010 Sep;18(9): 1640-1649.

Zhao H, Pestina TI, Nasimuzzaman M, Mehta P, Hargrove PW, Persons DA (2009) Amelioration of murine beta-thalassemia through drug selection of hematopoietic stem cells transduced with a lentiviral vector encoding both gamma-globin and the MGMT drug-resistance gene. Blood. Jun 4;113(23): 5747-5756.

Zhao W, Manley JLD (1998) Regulation of poly(A) polymerase interferes with cell growth. Mol Cell Biol. Sep;18(9): 5010-5020.

Zufferey R, Donello JE, Trono D, Hope TJ. (1999) Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. J Virol. Apr;73(4): 2886-2892.

Zychlinski D, Schambach A, Modlich U, Maetzig T, Meyer J, Grassman E, Mishra A, Baum C. (2008) Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol Ther. Apr;16(4): 718-725.

http://gassama.myweb.uga.edu/

http://www.alphathalassemia.net

http://envmedical.com

http://pathmicro.med.sc.edu/lecture/hiv9.htm

*http://*www.addgene.org/)